

CHAPITRE D

LES PRODUITS DU RECENSEMENT

***Remarque** : dans tout ce chapitre, on parlera des « résultats » du recensement et non pas des « données », de manière à réserver ce dernier mot à la désignation des informations élémentaires (qui ne sont pas communicables parce que couvertes par le secret statistique). L'agrégation des informations élémentaires conduit aux résultats mis à la disposition des utilisateurs dès lors qu'ils satisfont toutes les exigences de protection des données individuelles et de protection du secret statistique.*

D.1 Quelques précisions préalables

D.1.1 Multiplicité des utilisations des résultats du recensement

L'objet du recensement est de déterminer la population légale de toutes les collectivités territoriales et de toutes les circonscriptions administratives et de décrire les caractéristiques démographiques et sociales de la population et les caractéristiques des logements à tous les niveaux géographiques du territoire, du quartier (IRIS) dans les villes jusqu'au pays dans son ensemble.

Le recueil simultané d'informations portant sur la population et les logements offre de multiples possibilités, comme par exemple de mettre en rapport les caractéristiques des personnes avec celles de leur logement. De plus, par le recueil de données de localisation comme le lieu de travail ou le lieu de résidence des personnes cinq ans auparavant, le recensement peut fournir une description « au lieu de travail » des actifs travaillant dans une commune ou un territoire ou « au lieu de résidence antérieure » des personnes ayant quitté la commune ou la zone concernée.

Les utilisateurs de l'information issue du recensement sont légion.

Aux **communes**, le recensement fournit tout d'abord l'effectif officiel de la population, dont l'importance est primordiale de par ses implications, en termes financiers notamment. Au-delà de l'effectif, les collectivités locales sont grandement intéressées par la dynamique démographique et la mesure des flux migratoires, la pyramide des âges, le taux de chômage ou le taux de nouveaux arrivants. A cette connaissance générale au niveau de la commune s'ajoute une préoccupation de suivi des quartiers dans le cadre de la gestion infracommunale, sur des territoires dont le périmètre varie sensiblement selon les utilisations (*périmètre scolaire, transports, services de proximité, etc.*).

Par ailleurs, les communes sont de plus en plus impliquées dans des **structures intercommunales**, ce qui nécessite des informations homogènes et comparables entre les différentes communes et leurs quartiers. A cet égard, les **agences d'urbanisme** sont

particulièrement concernées par les problématiques d'agglomération et de déplacements urbains, mais aussi dans le cadre plus large des aires urbaines.

Les **collectivités départementales** utilisent beaucoup les statistiques relatives à la population, à sa répartition géographique et à ses caractéristiques démographiques et sociales pour les besoins de leur gestion et l'exercice de leur fonction de redistribution sociale.

Les **collectivités régionales** sont pour leur part intéressées par les migrations résidentielles interrégionales par sexe et âge, ou à défaut par des descriptions des flux de migration sur des zones plus larges. Les migrations liées à l'emploi ou à la retraite constituent pour elles des enjeux importants. Elles ont besoin de statistiques sur l'emploi et la formation, domaines dans lesquels elles ont des compétences étendues. Leurs préoccupations ne se limitent pas à l'échelon régional mais portent également sur les niveaux infrarégionaux.

Les **chercheurs** ont, eux aussi, des attentes multiples. Ils s'intéressent en particulier à la comparabilité des résultats dans l'espace et dans le temps, aux aspects géographiques et à la possibilité de reconstituer des agrégats, ainsi qu'à la continuité de l'information entre le nouveau recensement et les précédents. Enfin, l'accès aux variables sensibles⁵³ est un sujet de préoccupation des chercheurs, pour qui elles constituent un thème important d'étude.

Les **administrations** utilisent largement les résultats du recensement en fonction de leurs domaines de compétence : territoires ruraux ou urbains, gestion de l'eau, répartition des hommes et des activités sur les territoires, politique de la ville, évaluation des politiques publiques, etc. Toutes expriment le besoin de disposer sur l'ensemble du territoire d'informations homogènes, structurelles et en évolution, et en particulier de résultats sur les migrations ou les déplacements domicile-travail.

Enfin, des **sociétés privées** réalisent à partir des résultats détaillés du recensement, le plus souvent au niveau infracommunal, des produits qui s'adressent à différents types de clients : des entreprises, qui utilisent ces informations à des fins de marketing ou de constitution de dossiers administratifs (cas des implantations commerciales), mais également des collectivités locales. Elles expriment un fort besoin d'actualisation des résultats les plus fins, mais aussi de mise à disposition par l'Insee de bases de données de format adapté et à faible coût dans le domaine du "géomarketing".

La multiplicité des besoins qui s'expriment, la variété des utilisations et le grand nombre des utilisateurs ont pour conséquence que le recensement de la population demeure en France – pays qui a fait le choix de ne pas mettre en place des registres de population exhaustifs et interconnectables – une source statistique irremplaçable aux niveaux communal et infracommunal. Ce succès est naturellement porteur d'insatisfactions puisque certains des besoins exprimés sont antagoniques. Ainsi, sauf à faire tous les ans un recensement général de la population, ce qui serait d'un coût exorbitant et serait au surplus inutile puisque les structures démographiques n'évoluent pas très vite, il n'est pas possible d'obtenir une information qui, à la fois, comporterait une observation exhaustive des plus petites portions du territoire et serait à tout instant de la plus grande fraîcheur.

Dans ce contexte, le choix essentiel dans la rénovation du recensement de la population a été de mettre à la disposition des utilisateurs, qui tous ont besoin d'informations de référence sur les ménages, une information à la fois récente et fiable.

⁵³ Les données sensibles sont celles qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci.

D.1.2 Les deux grandes catégories de résultats du recensement de la population

Le dénombrement de la population de chaque collectivité territoriale et de chaque circonscription administrative du pays constitue l'objectif en quelque sorte régalien du recensement. Confié à l'Insee par la loi, ce dénombrement est authentifié par décret, ce qui revient à dire qu'il revêt un caractère officiel et s'impose pour l'application des multiples textes qui utilisent le chiffre de population pour la détermination d'un droit, par exemple le montant d'une dotation financière à une commune ou le nombre des membres du conseil municipal. Le vocabulaire courant parle de **population légale** du fait qu'elle est authentifiée par décret. La population légale ne se prête à aucune approximation. Certes, elle peut être contestée devant le juge administratif (qui aura alors à apprécier si l'Insee l'a établie au mieux compte tenu des informations disponibles qu'il était habilité à utiliser), mais elle ne saurait être présentée comme entourée d'un halo d'incertitude puisque cela rendrait discutable toute décision s'y référant.

La description des caractéristiques de la population et des logements qu'elle occupe relève d'une autre logique. On est cette fois dans l'univers de la production statistique et on parlera donc de **résultats statistiques**, éventuellement assortis d'une marge d'incertitude et par conséquent d'un intervalle de confiance. A moins qu'un texte juridique n'impose l'utilisation de telle ou telle statistique dans la détermination d'un droit, l'utilisateur est libre de retenir la statistique issue du recensement ou de l'écartier au profit d'une autre qu'il estimerait plus pertinente pour son propos.

La distinction est importante. Elle justifie, par exemple, que la somme des populations statistiques des quartiers d'une commune ne soit pas nécessairement égale à la population légale de cette commune. En effet, dans le nouveau recensement comme dans le recensement général de la population, certaines variables (l'activité, la profession détaillée, par exemple) ne sont exploitées, pour des raisons de coût, que pour une fraction de la population. L'extrapolation des échantillons conduit alors à des chiffres qu'il serait artificiel de vouloir corriger de manière que leur somme redonne exactement la population légale.

La suite du chapitre est donc consacrée à la présentation des calculs de détermination de la population légale (§ [D.2](#) et [D.3](#)), puis à celle des résultats statistiques du recensement (§ [D.4](#) et [D.5](#)). Enfin, le paragraphe [D.6](#) décrit la montée en régime de croisière de la production des résultats statistiques pendant la période 2004-2008.

D.2 Détermination de la population légale des communes

D.2.1 Aspects juridiques

La détermination des chiffres de population légale (population municipale, population comptée à part et population totale, dont les définitions sont données ci-dessous au [§ D.2.1.2](#)) est le premier objectif du recensement, y compris dans sa version renouvelée. La loi n° 2002-276 relative à la démocratie de proximité comporte dans son article 156, au paragraphe VII, les dispositions suivantes⁵⁴ :

VII. - Pour établir les chiffres de la population, l'Institut national de la statistique et des études économiques utilise les informations collectées dans chaque commune au moyen d'enquêtes de recensement exhaustives ou par sondage, les données démographiques non

⁵⁴ Le texte du titre V de la loi n° 2002-276 du 27 février 2002 figure en [annexe E.1](#).

nominatives issues des fichiers administratifs, notamment sociaux et fiscaux, que l'institut est habilité à collecter à des fins exclusivement statistiques, ainsi que les résultats de toutes autres enquêtes statistiques réalisées en application de l'article 2 de la loi n° 51-711 du 7 juin 1951 précitée.

Par ailleurs, il est rappelé que les communes de moins de 10 000 habitants sont recensées exhaustivement tous les cinq ans par roulement et que les communes de 10 000 habitants ou plus font l'objet d'une enquête par sondage tous les ans, 40% de leur population étant recensée au bout de cinq ans.

D.2.1.1 La périodicité de publication

L'article 156 de la loi prévoit, au paragraphe VIII, les dispositions suivantes :

VIII. - Un décret authentifie chaque année les chiffres des populations de métropole, des départements d'outre-mer et de Saint-Pierre-et-Miquelon, des circonscriptions administratives et des collectivités territoriales.

Le principe est donc de publier tous les ans les chiffres des populations de toutes les collectivités et circonscriptions. Pour les établir, l'Insee dispose des informations collectées à l'occasion des enquêtes de recensement et de données non nominatives issues de sources administratives. Ces sources administratives vont être utilisées pour actualiser les chiffres de population légale des communes de moins de 10 000 habitants entre deux enquêtes de recensement.

La détermination de la population légale suppose qu'on ait exploité les données collectées en début d'année, c'est-à-dire qu'elles aient été saisies et contrôlées. En fait, le décret est publié avant la fin de l'année pour prendre effet le 1er janvier de l'année suivante.

En Nouvelle-Calédonie, en Polynésie française, dans les Iles Wallis et Futuna et à Mayotte, la population des circonscriptions administratives est établie à l'issue des recensements généraux de la population, une fois tous les cinq ans.

D.2.1.2 Les catégories de population

Le décret n° 2003-485 du 5 juin 2003 fixe les catégories de population et leur composition⁵⁵.

La **population municipale** comprend les personnes ayant leur résidence habituelle sur le territoire de la commune, les personnes détenues dans les établissements pénitentiaires de la commune, les personnes sans abri recensées sur le territoire de la commune et les personnes résidant habituellement dans des habitations mobiles, recensées sur le territoire de la commune. L'enquête de recensement sur le territoire de la commune permet de déterminer cette population.

La **population comptée à part** comprend certaines personnes dont la résidence habituelle est dans une autre commune mais qui ont conservé une résidence sur le territoire de la commune :

- personnes mineures dont la résidence familiale est dans une autre commune mais qui résident du fait de leurs études dans la commune⁵⁶ ;

⁵⁵ Ce texte figure en [annexe E.5](#).

⁵⁶ Si ces personnes résident dans un logement de la commune, elles sont déclarées dans la liste C de la feuille de logement ; si elles résident dans un internat de la commune, elles sont repérées à partir

- personnes résidant dans une communauté de catégorie 1, 2 ou 3 (services de moyen ou de long séjour des établissements publics ou privés de santé, établissements sociaux de moyen ou long séjour, maisons de retraite, foyers et résidences sociales ; communautés religieuses ; casernes ou établissements militaires) d'une autre commune mais ayant une résidence familiale sur le territoire de la commune : elles précisent l'adresse de leur résidence personnelle dans le bulletin individuel recueilli dans leur communauté⁵⁷ ;
- personnes majeures âgées de moins de 25 ans ayant leur résidence familiale sur le territoire de la commune et qui résident dans une autre commune pour leurs études : elles sont déclarées dans la liste B de la feuille de logement de leur résidence familiale ;
- personnes sans domicile fixe rattachées à la commune au sens de la loi du 3 janvier 1969 et non recensées dans la commune : elles sont connues par les listes préfectorales.

Exemples :

M. et Mme DUPONT habitent ORLEANS. Ils ont trois enfants.

M. DUPONT père vit pendant la semaine à PARIS pour des raisons professionnelles et rentre à ORLEANS tous les week-ends.

Il sera compté dans la population municipale d'ORLEANS, sa résidence habituelle étant d'après les textes⁵⁸ sa résidence familiale.

M. et Mme DUPONT ont une fille de 20 ans, Alice, qui est en internat à BLOIS. Elle rentre chez ses parents tous les week-ends à ORLEANS.

Etant majeure de moins de 25 ans, elle sera comptée dans la population comptée à part d'ORLEANS où elle a sa résidence familiale et dans la population municipale à BLOIS où se trouve son internat, sa résidence habituelle.

Leur fils de 16 ans, François, fait ses études à TOURS où il habite chez son oncle. Il rentre à ORLEANS tous les week-ends.

Etant mineur, François DUPONT sera compté dans la population comptée à part de la ville de TOURS, ville où il réside pour ses études, et dans la population municipale de la ville d'ORLEANS, où se trouve la résidence de sa famille.

M. et Mme DUPONT ont aussi un autre fils, Damien, de 26 ans qui poursuit ses études à PARIS. Il a un studio et revient à ORLEANS tous les week-ends.

Il sera compté dans la population municipale à PARIS, lieu de sa résidence habituelle.

La mère de Mme DUPONT est en maison de retraite à TOURS mais elle a conservé sa résidence familiale à ORLEANS.

Elle sera comptée dans la population municipale de TOURS, ville où elle réside habituellement en maison de retraite, et dans la population comptée à part d'ORLEANS, où elle a sa résidence familiale.

La **population totale** est la somme de la population municipale et de la population comptée à part.

de la déclaration, dans leur bulletin individuel recueilli dans l'internat, d'une adresse de résidence familiale et de leur date de naissance.

⁵⁷ Les deux communes n'étant peut-être pas recensées en même temps, les effectifs des flux entre les communautés d'une commune et les autres communes doivent être conservés pendant cinq ans.

⁵⁸ Décret n° 2003-485 du 5 juin 2003, article 1§ III.

Un graphique récapitulatif présente de manière visuelle en [annexe D.1](#) la composition des catégories de population légale.

D.2.2 Une même date pour les résultats de toutes les communes

La méthode de calcul retenue consiste à produire chaque année, pour chaque commune, une population prenant effet juridique le 1er janvier suivant mais calculée en se référant à l'année médiane des cinq années écoulées.

Sur le plan technique, cette référence à l'année du milieu des cinq années écoulées tient au fait que les communes de 10 000 habitants ou plus sont enquêtées par sondage à raison de 40% de leur population en cinq ans, l'ensemble de leur territoire étant pris en compte au bout de ces cinq années, et que les communes de moins de 10 000 habitants sont recensées exhaustivement par roulement tous les cinq ans. L'ensemble du territoire national (métropole, départements d'outre-mer et Saint-Pierre-et-Miquelon) est ainsi couvert en cinq ans.

Cette option assure l'égalité de traitement des communes. En effet, il ne serait pas envisageable de publier des populations des communes qui ne seraient pas relatives à une même date, notamment pour les années de consultation électorale. Le Conseil d'Etat, consulté en 1998, a mis ce point fortement en avant (voir chapitre E, [§ E.1](#)).

Elle assure aussi une plus grande équité dans la répartition des concours de l'Etat, puisqu'elle confère aux dénombremens de population, qui sont le critère essentiel de cette répartition, le maximum de fiabilité en l'état actuel des techniques d'estimation⁵⁹. Les événements affectant, positivement ou négativement, la démographie d'une commune seront pris en compte avec un décalage de deux ans dans les communes de moins de 10 000 habitants et le seront par cinquièmes annuels successifs à partir de l'année suivant celle de leur déclenchement dans les communes de 10 000 habitants ou plus. Rappelons que de tels événements n'étaient pas pris en compte avant huit ou neuf ans quand ils se produisaient juste après un recensement général de la population.

L'Insee ne travaille pas en solitaire sur ces méthodes de calcul. Le Conseil national de l'information statistique (CNIS) donnera, tous les cinq ans, un avis quant au mode de détermination de la population légale des communes. Cet avis s'appuiera sur une évaluation des conditions de production des populations légales, sur une large concertation visant à recueillir le point de vue des utilisateurs de ces populations (administrations et collectivités locales) ainsi que sur un état des lieux des différentes sources administratives, de façon à apprécier les possibilités de progrès apportées par ces données⁶⁰.

⁵⁹ Cette plus grande fiabilité résulte des considérations suivantes. Ramener les populations à la quatrième ou à la dernière année de la période de cinq ans faisant l'objet des calculs obligerait, pour les communes de moins de 10 000 habitants, à extrapoler sur trois ou quatre ans les données les plus anciennes. Dans celles de 10 000 habitants ou plus, cela obligerait à introduire dans les calculs des coefficients correcteurs des tendances de croissance ou de décroissance de la population. Dans les deux cas, les dénombremens y perdraient en robustesse et le système y perdrait en lisibilité. Au surplus, les sources administratives ne sont pas disponibles pour l'année en cours, ce qui rendrait l'extrapolation véritablement hasardeuse la dernière année dans l'univers des communes de moins de 10 000 habitants.

⁶⁰ On peut imaginer de faire évoluer le système à l'avenir, une fois qu'il aura été rodé et que des progrès substantiels auront été réalisés en matière de fichiers administratifs, tant en ce qui concerne leur couverture que leur date de disponibilité. A la condition expresse que les calculs n'y perdent pas en sûreté, une première étape consisterait à ramener toutes les estimations à l'avant-dernière année de la période de cinq ans prise en compte, l'étape ultime de progrès visant l'année de l'enquête de recensement la plus récente. C'est un objectif pour la décennie 2010, qui doit être abordé avec prudence compte tenu des arguments exposés à la note précédente de bas de page.

D.3 Méthodes de calcul de la population légale

D.3.1 Méthode pour les communes de 10 000 habitants ou plus

D.3.1.1 Population des ménages

Les adresses des communes de 10 000 habitants ou plus sont réparties en cinq groupes. Chaque année, on enquête environ 40% des logements d'un groupe, ce qui représente 8% des logements de la commune.

En multipliant l'effectif de la population ainsi recensée par le poids de sondage, soit environ 2,5 (c'est le quotient du nombre de logements du groupe d'adresses par le nombre de logements effectivement recensés), on obtient une estimation de la population du groupe d'adresses.

En fin d'année N, les résultats des collectes effectuées en N-4, N-3, N-2, N-1 et N sont disponibles. On additionne les résultats ainsi obtenus au cours des cinq dernières années pour obtenir une estimation de la population moyenne et du nombre moyen de logements de la commune au cours de ces mêmes cinq dernières années.

Mais, ni la population moyenne des cinq années, ni le nombre moyen de logements, ne sont égaux à la population ou au nombre de logements de l'année médiane (N-2). Pour obtenir la population de l'année médiane, on fait l'hypothèse (extrêmement robuste) que le nombre moyen de personnes par logement calculé à partir des cinq collectes prises en compte est très voisin de celui de l'année médiane. Quant au nombre de logements de l'année médiane, c'est bien entendu celui que fournit le RIL.

Dès lors, pour obtenir la population de l'année médiane, il suffit de multiplier le nombre moyen de personnes par logement au cours des cinq années considérées par le nombre de logements de l'année médiane issu du RIL. On se situera au 1er janvier de cette année médiane. Un exemple chiffré est donné ci-après.

Remarques :

1- Afin de tenir compte de la spécificité du parc des logements de certaines communes, le plan de sondage a été conçu de façon à traiter exhaustivement les "adresses de grande taille", c'est-à-dire celles qui, à la fois, sont les plus grandes, comptent chacune au moins 60 logements et représentent 10% des logements de la commune⁶¹ ; par ailleurs, les adresses nouvelles de toutes ces communes sont toutes enquêtées dans les cinq années de leur apparition ; ainsi, dans les communes de 10 000 habitants ou plus, le calcul est fait d'une part pour les adresses de grande taille, d'autre part pour les adresses nouvelles, enfin pour le reste des logements⁶².

2- La méthode s'applique aussi dans les DOM, en tenant compte de la spécificité du plan de sondage⁶³.

⁶¹ Ces critères ont été arrêtés au terme de très nombreuses simulations destinées à rechercher la meilleure optimisation du plan de sondage.

⁶² Voir chapitre B, [§ B.3.2](#)

⁶³ Voir chapitre B, [§ B.3.2, question 9](#).

Exemple chiffré

Année de publication de la population légale	2015	2016	2017	2018	2019	2020
Année d' estimation de la population des ménages	2013	2014	2015	2016	2017	2018
Nombre total de logements (L) dans le répertoire d'immeubles localisés (RIL) de la commune l'année d'estimation	6 000	6 100	6 150	6 200	6 230	6 300
Nombre de logements du groupe de rotation d'adresses de l'année d'estimation	1 200	1 250	1 300	1 180	1 200	1 260
Population estimée du groupe d'adresses	2 400	2 500	2 600	2 360	2 400	2 520
Somme des populations des 5 groupes d'adresses (p)			12 260	12 380		
Somme des logements des 5 groupes d'adresses (l)			6 130	6 190		
Nombre moyen observé de personnes par logement (p/l)			$\frac{12\ 260}{6\ 130}$	$\frac{12\ 380}{6\ 190}$		
Population (L x p/l)			12 300	12 400		

D.3.1.2 Population des communautés

La population des communautés est recensée exhaustivement (par l'Insee) une fois tous les cinq ans dans les communes de 10 000 habitants ou plus. Celles-ci ont été réparties en cinq groupes de manière à étaler la charge sur le cycle de cinq ans. Une exception est faite pour les communes comportant un nombre très élevé de communautés : pour alléger la charge de la collecte dans ces communes, le recensement des communautés y est réparti sur plusieurs années.

En fin d'année N, pour une commune donnée, il s'agit d'estimer la population de ses communautés au 01/01/N-2, à la même date que la population des ménages.

Si les communautés de la commune ont été recensées en N-2, on retient le chiffre issu de la collecte.

Si elles ont été recensées avant, en N-3 ou N-4, on actualise la population de leurs communautés grâce au répertoire des communautés⁶⁴. L'extrapolation consistera à faire évoluer la population recensée en communauté en prenant en compte les évolutions observées dans le répertoire : on ajoute la population des nouvelles communautés et on retranche celle des communautés disparues.

Si les communautés de la commune ont été recensées après, en N-1 ou N, on actualise la population de leurs communautés par interpolation entre le dernier chiffre publié et celui fourni par le recensement des communautés.

NB : pour les communes dans lesquelles le recensement des communautés est réparti sur plusieurs années, la population de chaque sous-ensemble est estimée séparément, selon la méthode exposée ci-dessus, en fonction de l'année de collecte de chacun des sous-ensembles.

⁶⁴ Voir chapitre C, § C.3.

D.3.1.3 Population des habitations mobiles, personnes sans abri, mariniers

Les personnes résidant habituellement dans des habitations mobiles terrestres et les personnes sans abri sont recensées exhaustivement par les agents recenseurs les deux premiers jours de la collecte faite auprès des ménages, la situation enregistrée étant celle existant sur le terrain le premier jour de la collecte à zéro heure⁶⁵. Les effectifs ainsi dénombrés sont maintenus constants pendant cinq ans et mis à jour après chaque nouvelle enquête exhaustive de cette nature. Il en est de même pour les mariniers, qui sont recensés par l'Insee.

D.3.1.4 Population comptée à part

Elle est mise à jour tous les ans au fur et à mesure des enquêtes. Par exemple, les effectifs de personnes résidant en communauté dans d'autres communes et ayant une résidence personnelle dans la commune sont actualisés au fur et à mesure du recensement de ces communes.

D.3.2 Méthode pour les communes de moins de 10 000 habitants

Ces communes sont recensées exhaustivement, par roulement, tous les cinq ans, à raison d'un cinquième des communes chaque année.

D.3.2.1 Population des ménages

Pour la suite du propos, on se situe en fin d'année N.

Sachant que la population publiée pour les communes de 10 000 habitants ou plus sera celle relative à l'année médiane des cinq années écoulées, soit N-2, c'est la population au 01/01/N-2 qu'il s'agit d'estimer pour chacune des communes de moins de 10 000 habitants.

Pour les communes recensées en N-2, on retient le résultat de l'enquête de recensement. Pour les autres, il convient d'actualiser la population de la commune publiée l'année précédente.

Pour celles qui ont été recensées en N-1 ou N, on procède à cette actualisation par interpolation avec le dernier chiffre publié (relatif à N-3). Cela consiste à calculer l'écart entre ce chiffre et celui issu de l'enquête de recensement de N-1 ou N et à le répartir à parts égales sur les deux ou les trois années.

Pour celles qui ont été recensées en N-3, on actualise le chiffre tiré de l'enquête de recensement N-3 en s'aidant d'une information supplémentaire sur l'évolution du parc de logements : cette information est tirée du fichier de la taxe d'habitation et le calcul est fait pour chaque commune.

Il faut alors relier l'évolution du parc de logements et celle du nombre d'habitants et tenir compte du fait que, souvent, la population et le nombre de ménages ne croissent pas exactement au même rythme. En moyenne, et dans la plupart des communes, la taille des ménages diminuant, le nombre de ménages augmente plus vite que celui des habitants.

⁶⁵ Voir au chapitre C, [§ C.2.2.2](#), les modalités de la collecte.

Pour passer d'une évolution du parc de logements à une évolution du nombre d'habitants, on estime, à l'aide des dernières enquêtes de recensement disponibles, la différence entre les deux rythmes d'évolution. On reporte ensuite cette différence sur le taux d'évolution du parc calculé au moyen des fichiers de la taxe d'habitation afin d'avoir une estimation du taux de progression de la population.

Admettons qu'entre les deux derniers recensements le taux annuel d'évolution de la population de la commune Z a été de 1,5% pour les logements et de 1,3% pour les personnes. On retient que le parc de logements de Z croît plus vite (de 0,2 point) que le nombre d'habitants. Si les données du fichier de la taxe d'habitation font apparaître, pour la commune Z et pour l'année considérée, une augmentation du parc de logements de 1,6%, l'évolution de la population de Z sera estimée à $1,6 - 0,2 = 1,4\%$.

Pour les communes recensées en N-4, on procède de la même façon, en partant du dernier chiffre de population publié, c'est-à-dire celui du 01/01/N-3. Cela correspond à une extrapolation sur deux ans du résultat de l'enquête de recensement.

Au cours du cycle quinquennal glissant, chaque commune passe donc par les cinq " états " suivants : interpolation sur deux ans, interpolation sur un an, résultat de l'enquête de recensement, extrapolation sur un an, extrapolation sur deux ans. Le chiffre de population est toujours calculé sur la base d'une enquête de recensement distante d'au plus deux ans.

Exemple chiffré

Considérons la commune Z, recensée en 2006 et 2011.

Le recensement de 1999 donnait 1 600 habitants et 600 logements.

L'enquête de recensement de 2006 donne 2 000 habitants et 800 logements.

Le fichier de la taxe d'habitation (TH) donne 800 logements en 2006, 830 en 2007, 850 en 2008 et 870 en 2009.

Fin 2008, la population retenue sera celle de l'enquête de recensement de 2006, c'est-à-dire 2 000 habitants.

Fin 2009, il s'agit de calculer une population au 01/01/2007.

L'évolution du parc de logements est estimée à 3,75 % entre le 01/01/06 et le 01/01/07 (passage de 800 à 830 logements). Entre les deux derniers recensements connus (en l'occurrence 1999 et 2006), on a observé que le nombre de logements n'avait pas crû au même rythme que le nombre d'habitants (4,2% par an pour le nombre de logements, 3,2% par an pour le nombre d'habitants). En reportant cette différence, on estime la progression du nombre d'habitants à 2,75% (un point de moins que la progression du nombre de logements). La population estimée au 01/01/07 sera donc de $2\ 000 \times 1,0275$, soit 2 055 habitants.

Fin 2010, il s'agit d'actualiser la population au 01/01/2008

L'évolution du parc de logements est estimée à 2,4 % entre le 01/01/07 et le 01/01/08 (passage de 830 à 850 logements). On retient donc une évolution du nombre d'habitants de 1,4 %. La population au 01/01/08 est ainsi estimée à $2\ 055 \times 1,014$.

Fin 2011, il s'agit d'actualiser la population au 01/01/2009.

On dispose du résultat de l'enquête de recensement de 2011, qui donne une population de 2 200 habitants. Cela représente une progression de 116 habitants par rapport à 2008. L'interpolation consiste à répartir cette progression à parts égales sur chaque année, soit 39 habitants par an.

La population au 01/01/09 est donc estimée à 2 123 habitants.

Fin 2012, il s'agit d'actualiser la population au 01/01/2010.

On poursuit l'interpolation, en ajoutant à nouveau 39 habitants. La population estimée au 01/01/10 est donc de 2 162 habitants.

Fin 2013, il s'agit d'actualiser la population au 01/01/2011.
On retient alors le chiffre du recensement de 2011, soit 2 200 habitants.

Organisation des calculs dans le temps pour la population des ménages

Plaçons-nous l'année A.

A la fin de l'année A, il convient de calculer la population légale estimée au début de l'année A-2.

Selon le groupe de rotation des communes auquel elles appartiennent, les communes de moins de 10 000 habitants ont été recensées l'une des cinq dernières années. Comment s'y prend-on pour estimer leur population au 1/1/A-2 ? Selon le cas (cf. tableau ci-après), on actualise par extrapolation si la dernière enquête est antérieure à la date d'estimation, ou par interpolation si la dernière enquête est postérieure à la date d'estimation. Dans tous les cas, il s'agit d'actualiser au 1/1/A-2 l'enquête de recensement la plus proche :

Groupe de rotation auquel appartient la commune	Année du recensement selon le groupe de rotation						
	A-6	A-5	A-4	A-3	A-2 (année de l'estimation)	A-1	A (année de la publication)
I			Recensement		Extrapoler de A-4 à A-2		
II				Recensement	Extrapoler de A-3 à A-2		
III					Utiliser le recensement qui a lieu cette année		
IV	Recensement				Interpoler entre A-1 et l'estimation A-3)	Recensement	
V		Recensement			Interpoler entre A et l'estimation A-3)		Recensement

Les deux années d'extrapolation pourraient, si l'on se contentait de prolonger les tendances, être hasardeuses dans le cas où des événements lourds de conséquences démographiques se produiraient juste après le recensement.

C'est pour cela que les extrapolations sont construites en recourant aux informations auxiliaires contenues dans les fichiers de la taxe d'habitation (TH). En A, on dispose en effet des fichiers TH des années A-1 et précédentes. Or, une analyse approfondie de la source TH a montré qu'elle permettait d'estimer les taux d'évolution du nombre de résidences principales avec une bonne précision dans la plupart des communes.

En revanche, les deux années d'interpolation ne présentent guère de risque de s'écarter de la réalité puisqu'on s'appuie alors sur une enquête de recensement exhaustive toute récente. Le recours aux fichiers administratifs est ici superflu.

Pour raccorder les extrapolations et les interpolations entre deux enquêtes de recensement et ainsi lisser les évolutions, on fait l'interpolation entre la dernière extrapolation et la nouvelle enquête de recensement. Ainsi, dans le groupe V, qui a connu une enquête de

recensement début A-5 et début A, on interpole, non pas entre ces deux enquêtes, mais entre les résultats du recensement A et la dernière extrapolation conduite à partir de A-5, qui a donné l'estimation de l'année A-3.

D.3.2.2 Population des communautés

Le principe est identique à celui retenu pour la population des ménages, à savoir extrapolation et interpolation.

Il est rappelé que le recensement des communautés est opéré (par l'Insee) de manière exhaustive et en une seule fois tous les cinq ans sur le territoire de la commune⁶⁶.

Le principe d'extrapolation est le même que pour la population des ménages, c'est-à-dire qu'il concerne les deux années postérieures au recensement dans la commune. La source utilisée est ici le répertoire des communautés constitué par l'Insee. L'extrapolation consiste à faire évoluer la population recensée en communauté comme observé dans le répertoire : en ajoutant la population correspondant aux nouvelles communautés et en retranchant la population des communautés disparues.

Enfin, pour les deux années antérieures au dernier décompte, la population des communautés de la commune est calculée par interpolation entre les populations recensées aux deux derniers recensements.

D.3.2.3 Population des habitations mobiles, personnes sans abri, marinières

Les personnes résidant habituellement dans des habitations mobiles terrestres et les personnes sans abri sont recensées exhaustivement par les agents recenseurs les deux premiers jours de la collecte faite auprès des ménages, la situation enregistrée étant celle existant sur le terrain le premier jour de la collecte à zéro heure⁶⁷. Les effectifs ainsi dénombrés sont maintenus constants pendant cinq ans et mis à jour après chaque nouvelle enquête exhaustive de cette nature. Il en est de même pour les marinières, qui sont recensés par l'Insee.

D.3.2.4 Population comptée à part

Elle est mise à jour tous les ans au fur et à mesure des enquêtes. Par exemple, les effectifs de personnes résidant en communauté dans d'autres communes et ayant une résidence personnelle dans la commune sont actualisés au fur et à mesure du recensement de ces communes.

D.3.3 La population des fractions cantonales et des communes associées

La population municipale de la commune est répartie dans les fractions cantonales ou les communes associées proportionnellement aux dénombrements exhaustifs dont on dispose. Dans le cas d'une commune de moins de 10 000 habitants, il s'agit du dernier recensement la concernant ; pour les autres communes, il s'agit du nombre de logements.

⁶⁶ Exception faite pour les communes de 10 000 habitants ou plus dans lesquelles la charge a été répartie sur plusieurs années (voir chapitre B, [§ B.3.3.1](#)).

⁶⁷ Voir au chapitre B, [§ B.3.3.2 à B.3.3.4](#), les modalités de la collecte.

D.4 Les résultats statistiques

D.4.1 Deux catégories de résultats statistiques

Le recensement repose désormais sur des enquêtes répétées annuellement auprès de populations différentes. Ainsi, les communes de moins de 10 000 habitants sont réparties en cinq groupes, de même que les adresses des communes de 10 000 habitants ou plus. Les groupes de communes de moins de 10 000 habitants sont recensés successivement chaque année. Dans les communes de 10 000 habitants ou plus, les groupes d'adresses sont enquêtés successivement chaque année avec un taux de sondage de 40%. Au bout de cinq ans, tout le territoire a été pris en compte et l'ensemble de données ainsi constitué couvre la totalité des logements des communes de moins de 10 000 habitants et 40% des logements des communes de 10 000 habitants ou plus.

La construction des résultats statistiques du recensement va utiliser ces données de deux manières :

- à *tous les niveaux géographiques*⁶⁸, on utilise l'ensemble des données collectées au cours des cinq dernières enquêtes de recensement pour produire une information calée sur l'année médiane et rendant compte de la situation moyenne de la période quinquennale correspondante. Chaque année A, les données les plus anciennes (celles de l'année A-5) sont remplacées par les données collectées début A. L'ancienneté des données est donc toujours la même, de trois ans à quatre ans. En effet, la disponibilité des données intervient en fin d'année, et elles sont utilisées pendant un an avant d'être renouvelées⁶⁹ ;
- pour la *France et les régions*, le fait que les groupes de rotation des communes de moins de 10 000 habitants soient équilibrés au niveau régional et que les groupes de rotation d'adresses le soient au niveau de la commune va permettre de produire, avec les seules données collectées d'une année, des résultats significatifs aux niveaux national et régional. Les résultats auront dans ce cas une ancienneté de 1 à 2 ans, selon la date d'utilisation.

La première catégorie de résultats est dénommée « **résultats du recensement** » car elle constitue l'essentiel de la production du recensement. La seconde est appelée « **résultats de l'enquête annuelle de recensement** ».

Les développements qui suivent sont consacrés à la présentation des modes d'élaboration et de la précision de ces deux catégories de résultats, aux différents niveaux d'utilisation qui peuvent en être faits et aux dispositions particulières adoptées pendant la période de montée en régime de croisière (2004-2008).

Remarque :

⁶⁸ De la France à la commune, et même à l'IRIS-2000 pour les communes découpées en IRIS-2000 (îlots regroupés pour l'information statistique), qui constituent des quartiers d'environ 2 000 habitants.

⁶⁹ Selon les niveaux géographiques, il serait envisageable de produire des résultats à partir de deux, trois ou quatre années de collecte, afin d'accroître leur fraîcheur. C'est le choix du bureau du Census américain pour la diffusion de l'*American Community Survey*, qui remplace le questionnaire long du recensement des Etats-Unis. L'Insee a écarté cette manière de faire dans l'immédiat afin de ne pas compliquer l'utilisation des résultats du nouveau recensement. Il l'a cependant adoptée pour la période de montée en régime de croisière (2004-2008), afin de produire de premiers résultats pour les communes de 10 000 habitants ou plus (cf. [§ D.6 ci-dessous](#)).

Les utilisateurs de résultats des recensements généraux de la population se souviennent que l'exploitation des données n'était pas nécessairement exhaustive. Les variables les plus complexes, comme la profession détaillée, la structure des ménages, etc., étaient exploitées à partir d'un échantillon d'un quart de la base globale. La raison en est la forte charge de traitement que ces variables occasionnent, ce qui aurait conduit à un trop long délai de publication, accentuant encore le décalage avec la réalité, si on avait prétendu les exploiter toutes.

Avec le nouveau recensement, le principe du sondage au quart pour ces variables est conservé pour les communes de moins de 10 000 habitants. En revanche, pour les communes de 10 000 habitants ou plus, qui font déjà l'objet d'un sondage à 40 % en cinq ans, l'intégralité des données sera exploitée. Ce traitement différencié est justifié par le poids, dans ces communes, des problématiques liées à la profession et à l'activité économique.

D.4.2 Mode d'élaboration des résultats du recensement

D.4.2.1 Quelles données pour quels résultats ?

A partir de l'année 2008, qui marquera la réalisation de la cinquième enquête de recensement, c'est-à-dire l'achèvement de la prise en compte de la totalité du territoire, l'ensemble des données recueillies pourront être utilisées pour élaborer des descriptions statistiques à tous les niveaux, à partir du niveau IRIS-2000, niveau le plus fin pour la diffusion. Ces descriptions seront actualisées tous les ans à partir de 2008, faisant chaque fois masse de cinq années de collecte successives.

Concrètement, les données détaillées annuelles ont la forme d'un fichier de données individuelles, anonymisées et pondérées en fonction de leur échantillon d'origine (fichier-détail).

Ce fichier est tabulable pour toute analyse. Il rassemble les données de chaque commune et, pour celles de 10 000 habitants ou plus, de chaque IRIS-2000. Il permet des études sur ce territoire de base et sur tous les regroupements souhaités d'IRIS-2000.

En régime permanent, à la fin de l'année A, le fichier-détail incorpore les cinq dernières collectes, de A-4 à A. Le choix du jeu de pondérations est guidé avant tout par la volonté que les statistiques qui seront produites soient représentatives de l'ensemble de la période. Les pondérations sont donc centrées sur la date située au milieu de la période de collecte, c'est-à-dire le 1er janvier de l'année A-2.

C'est ainsi que le fichier construit à la fin de 2008 sera « millésimé » 01-01-2006, celui construit à la fin 2009 sera millésimé 01-01-2007, et ainsi de suite.

Il y a donc deux types d'opérations à mener, dont ce document expose le principe pour le fichier construit en 2008 et millésimé 01-01-2006 :

- celles qui consistent, pour les communes de moins de 10 000 habitants, recensées par roulement de 2004 à 2008, à amener toutes leurs données au 1er janvier 2006. Concrètement, il faut faire avancer dans le temps les données des communes recensées en 2004 et 2005, et ramener en arrière les données des communes recensées en 2007 et 2008 ;

- celles qui consistent, pour les communes de 10 000 habitants ou plus, à rassembler les échantillons enquêtés entre 2004 et 2008, qui portent au total sur 40% des logements de chacune de ces communes, et à les organiser pour produire la meilleure image « moyenne » au 1er janvier 2006.

En 2009, on renouvellera ces opérations en décalant d'un an, et ainsi de suite.

D.4.2.2 Les résultats relatifs aux communes de moins de 10 000 habitants

D.4.2.2.1 Méthode pour la population des résidences principales

Pour la population des résidences principales des communes de moins de 10 000 habitants, les données en niveau sont ajustées entre deux enquêtes de recensement successives grâce à des coefficients d'évolution résultant de l'exploitation de fichiers administratifs, mais les données en structure sont invariantes : ce sont celles issues de la plus proche enquête de recensement.

Pour une commune recensée en 2004, les structures alors observées de sa population (composition par sexe et âge, caractéristiques sociales, etc.), ne le seront à nouveau qu'à l'occasion de son prochain recensement, c'est-à-dire en 2009. Mais l'effectif de la population et des logements peut être extrapolé au 01-01-2006 en se servant de sources administratives complémentaires, notamment du fichier de la taxe d'habitation. A l'inverse, l'effectif de la population d'une commune recensée en 2008 peut être ramené, par interpolation entre deux collectes exhaustives, à la même date de référence. En régime courant, la démarche - la même que pour le calcul de la population légale - est la suivante.

Plaçons-nous l'année A. Selon le groupe de rotation auquel elles appartiennent, les communes de moins de 10 000 habitants ont été recensées l'une des cinq dernières années. Il s'agit donc d'estimer leur population au 1/1/A-2. Selon le cas (cf. tableau) on actualise par extrapolation si la dernière enquête de recensement est antérieure à la date d'estimation, ou par interpolation si la dernière enquête est postérieure à la date d'estimation. Dans tous les cas, il s'agit d'actualiser au 1/1/A-2 l'enquête de recensement la plus proche, selon le schéma suivant (dans lequel R désigne cette enquête quinquennale et exhaustive) :

Groupe de rotation	A-6	A-5	A-4	A-3	A-2 (année d'estimation)	A-1	A (année de publication)
I			R		Extrapoler A-4		
II				R	Extrapoler A-3		
III					R est utilisée		
IV	R				Interpoler entre A-6 et A-1 ⁷⁰	R	
V		R			Interpoler entre A-5 et A ⁷⁰		R

⁷⁰ On interpole en fait en partant de la dernière extrapolation calculée (ici A-3). Voir plus bas.

Puisqu'il tient compte de l'information la plus récente, le procédé d'interpolation entre deux enquêtes exhaustives a peu de risque de s'écarter de la réalité. En revanche, les deux années d'extrapolation pourraient, si l'on se contentait de prolonger les tendances, être hasardeuses si des événements affectant sensiblement la population se produisaient juste après le recensement.

Les extrapolations sont donc construites en recourant aux informations auxiliaires contenues dans les fichiers de la taxe d'habitation (TH). En A, on dispose en effet des fichiers TH des années A-1 et précédentes. Or, une analyse approfondie de la source TH a montré qu'elle permettait d'estimer les taux d'évolution du nombre de résidences principales avec une bonne précision dans la plupart des communes.

Il reste à tenir compte du différentiel entre le taux annuel d'évolution du parc des résidences principales et celui de la population. On a vérifié, aux recensements précédents, que dans une très grande majorité des cas, le premier est supérieur au second. Ne pas tenir compte de ce différentiel ferait que l'extrapolation viserait systématiquement trop haut. Soit d la différence entre le taux annuel d'évolution du parc de résidences principales (t) et le taux annuel d'évolution de la population observé pour une commune entre les deux recensements précédents ; le taux corrigé appliqué sera égal à :

$$t_{corr} = t - d$$

Par exemple si, pour une commune donnée, les taux d'évolution annuels constatés entre les deux derniers recensements sont de 5% pour les logements et 3% pour les personnes, le différentiel est de 2 points de pourcentage. Ce taux différentiel d'évolution permet de corriger le taux issu du fichier de la taxe d'habitation et de calculer le taux d'évolution de la population.

Enfin, pour raccorder les extrapolations et les interpolations entre deux enquêtes de recensement, on fait l'interpolation entre la dernière extrapolation et la nouvelle enquête.

Exemple chiffré

Soit une commune ayant 1 791 habitants à son recensement de 2013. A cette date, on constate que, depuis le recensement précédent, la différence d entre les taux annuels d'évolution des résidences principales et de la population correspondante est de 1,2%.

Dans les fichiers de la TH, on dénombre 694 résidences principales en 2013, 733 en 2014 (+5,6%) et 831 en 2015 (+13,4%).

Enfin, en 2018, le nouveau recensement dénombre 2 550 personnes.

Avec ces données la suite des populations sera :

Année de publication	2015	2016	2017	2018	2019	2020
Année d'estimation	2013	2014	2015	2016	2017	2018
Population	1 791 recense- ment	1 870 extrapola- tion	2 098 extrapola- tion	2 249 interpola- tion	2 400 interpola- tion	2 550 recense- ment

Explications : pour la première année d'extrapolation, on applique à la population du dernier recensement (1 791h.) le taux d'évolution : $t_{corr} = 5,6\% - 1,2\% = 4,4\%$. De la même façon, le taux d'évolution suivant sera de 12,2% (= 13,4 – 1,2). Pour les interpolations, l'écart entre les valeurs 2 098 (dernière estimation) et 2 550 (dernière enquête de recensement), à savoir 452, est réparti à parts égales sur trois années, soit 151 personnes. On en déduit aisément les populations de 2018 et 2019.

Le fichier-détail annuel des ménages

Restons dans la même commune. Pour l'année d'estimation 2013, le fichier-détail des ménages est exactement celui des données recueillies au recensement effectué cette année-là.

Pour l'année d'estimation 2014, on garde les mêmes données, mais on leur applique un coefficient qui tient compte de la variation estimée du niveau. En l'occurrence, le coefficient est égal à $1\,870 / 1\,791 = 1,044$.

Le même principe s'applique à l'année d'estimation 2015, le coefficient d'évolution associé étant alors égal à $1,171 (2\,098 / 1\,791)$.

Pour l'année d'estimation 2016, on dispose alors des données du nouveau recensement effectué en 2018. Le nouveau fichier est donc constitué de ces nouvelles données, auxquelles on rattache un coefficient pour ramener le niveau deux ans en arrière, c'est-à-dire $2\,249 / 2\,550 = 0,882$.

De la même façon, le coefficient appliqué pour l'année d'estimation 2017 vaut 0,941. Et ainsi de suite.

Quel est l'intérêt de ces coefficients ?

Ils sont à la disposition des utilisateurs pour l'analyse d'ensembles multicommunaux dans lesquels les communes ne sont pas nécessairement recensées toutes à la même date, par exemple un arrondissement. Dans ce cas, on utilise pour chaque commune de l'ensemble les données du plus proche recensement la concernant et on se sert des coefficients pour ajuster tous les effectifs à la même date de référence.

D.4.2.2 Méthode pour la population des communautés, des personnes résidant habituellement dans des habitations mobiles terrestres, des personnes sans abri et des marinières

Les communautés sont recensées en même temps que la population des ménages, c'est-à-dire une fois tous les cinq ans. Sous réserve de disposer de sources administratives de bonne qualité, on fait évoluer la population des communautés comme le taux d'évolution de ces sources administratives entre deux collectes. En A-2, les communautés des communes vues cette année-là compteront pour elles-mêmes et les communautés vues avant ou après la date d'estimation verront leurs données affectées d'un coefficient correcteur afin de ramener les données collectées à la date d'estimation.

Le fichier-détail sera un fichier d'individus agrémenté d'un coefficient s'appliquant à toutes les caractéristiques de chaque individu.

Pour les autres catégories de personnes, également recensées une fois tous les cinq ans, leurs effectifs et caractéristiques seront maintenus constants dans l'intervalle de deux enquêtes de recensement successives.

D.4.2.3 Les résultats relatifs aux communes de 10 000 habitants ou plus

D.4.2.3.1 Méthode pour la population des résidences principales

Les adresses des communes de 10 000 habitants ou plus sont réparties en cinq groupes de rotation. Chaque année, on enquête environ 40% des logements d'un groupe, ce qui représente 8% des logements de la commune (y compris les adresses neuves). En multipliant les données à l'adresse d'un groupe par leur poids de sondage (environ 2,5) tout

en respectant l'appartenance de l'adresse à sa strate - adresses de grande taille, adresses nouvelles et autres adresses - on obtient une estimation de la population et de ses caractéristiques au niveau du groupe d'adresses. On additionne les résultats des cinq collectes les plus récentes pour obtenir une estimation de la population moyenne de la commune et de ses caractéristiques au cours des cinq dernières années. Pour estimer la population l'année A-2, on « cale » ce résultat sur le nombre de logement du répertoire d'immeubles localisés (RIL) au début de l'année A-2. Concrètement, si le nombre moyen de logements de la commune est de 6 130, mais que le nombre de logements dans le RIL pour le début de l'année A-2 est de 6 150, on multipliera le poids initial de chaque adresse par le coefficient : $6\ 150/6\ 130$, soit 1,003.

Remarque :

Le plan de sondage inclut une stratification séparant les adresses de grande taille et les adresses nouvelles ; le calage s'opère donc par strate.

On obtient ainsi un fichier d'adresses dans lequel chaque adresse est assortie d'un poids s'appliquant à tous les logements et tous les individus de cette adresse.

D.4.2.3.2 Méthode pour la population des communautés, des personnes résidant habituellement dans des habitations terrestres mobiles, des personnes sans abri et des marinières

A l'inverse des ménages, les communautés d'une commune de 10 000 habitants ou plus ne font pas l'objet d'une enquête tous les ans. Sauf exception, elles sont recensées en une seule fois tous les cinq ans dans une même commune, les communes étant réparties dans le temps de manière à équilibrer la charge de collecte au niveau régional. Au bout de cinq ans, elles auront été toutes visitées. Comme pour les communes de moins de 10 000 habitants, sous réserve de disposer de sources administratives correctes, on fait évoluer la population de ces communautés comme le taux d'évolution de ces sources. L'année A-2, les communautés recensées cette année-là comptent pour elles-mêmes et l'on applique un coefficient correcteur aux données des communautés recensées avant et après la date d'estimation afin de ramener les données à la date d'estimation.

Les autres catégories de personnes (personnes en habitation mobile ou sans abri et marinières) sont beaucoup moins nombreuses. Elles sont également recensées une fois tous les cinq ans, mais des erreurs d'observation ont davantage de risque de se produire et on ne dispose pas de fichiers administratifs susceptibles d'aider à éliminer ces éventuelles erreurs et à mesurer des évolutions d'effectifs. Les effectifs et caractéristiques de ces personnes sont par conséquent maintenus constants dans l'intervalle de deux enquêtes de recensement successives.

D.4.2.4 Précision des résultats du recensement

La généralisation de la production de résultats à partir d'échantillons rend nécessaire un affichage clair de la précision des informations diffusées. C'est surtout au niveau infracommunal que le besoin d'information sur la fiabilité des données est le plus fort.

La mesure de la précision d'une statistique construite à partir d'un échantillon aléatoire (cas des communes de 10 000 habitants ou plus) s'effectue en considérant les deux types de risques pouvant affecter l'estimation, à savoir le biais et la dispersion.

Le biais mesure la capacité de l'estimation à être juste « en moyenne ». Par analogie avec un fusil, qui peut tirer systématiquement un peu trop à gauche ou à droite, un estimateur peut donner des résultats qui en moyenne sont sous-évalués ou sur-évalués. Ce biais peut être dû, par exemple, à un défaut d'exhaustivité de la base de sondage.

La dispersion mesure l'écart moyen autour de la moyenne (l'écart des impacts autour du centre de la cible pour reprendre l'analogie avec le fusil). Elle se mesure par la variance, qui se définit comme la moyenne des carrés des écarts à la moyenne, ou par l'écart-type qui en est la racine carrée.

Compte tenu du plan de sondage choisi, et en l'absence de défaut de la base de sondage, le biais des résultats du recensement est nul. La précision est donc représentée par la dispersion.

En l'absence de séries annuelles de population, les études sur la précision ont été menées par simulation. Dans un premier temps on a sélectionné des communes de tailles et caractéristiques d'habitat variables. Partant du fichier du recensement de 1999, on a simulé un grand nombre de fois des échantillons conformes au plan de sondage retenu. On a pu ainsi mesurer empiriquement la variabilité des données fournies pour une large gamme de variables observées.

Les utilisateurs ont souhaité un affichage de la précision en termes de coefficient de variation, qui donne une précision en pourcentage plutôt qu'un intervalle de confiance. Cependant, on passe aisément de l'un à l'autre. En effet, le coefficient de variation (CV) est le rapport de l'écart-type à la moyenne, tandis que l'intervalle de confiance à 95% est égal au résultat trouvé plus ou moins deux écarts-types. Il s'interprète de la façon suivante : pour un résultat donné, la « vraie » valeur a un peu plus de 95% de chances de se trouver dans l'intervalle compris entre le résultat moins deux écarts-types et le résultat plus deux écarts-types.

La précision dépend essentiellement de la taille de l'effectif étudié (c'est ainsi que la précision sur un ensemble de 5 000 personnes sera meilleure que sur un ensemble de 500)⁷¹. Voici l'échelle de précision qui peut être présentée au niveau de la commune (niveau pour lequel le plan de sondage est optimisé sur cinq ans)^{72, 73} pour des résultats moyens sur cinq ans :

Effectif	Coefficient de variation
< 500	> 6%
500 à <1 000	5%

⁷¹ De ce point de vue, il est intéressant de noter que, pour les variables d'activité économique et de profession détaillée, on dispose, dans les communes de 10 000 habitants ou plus, d'un échantillon de 40% sur cinq ans au lieu d'un échantillon instantané de 25% comme c'était le cas lors des recensements généraux.

⁷² Voir en [annexe D.2](#) la précision des résultats du recensement au niveau infracommunal.

⁷³ Le plan de sondage est optimal sur cinq ans parce que chacun des cinq échantillons utilisés pour construire les résultats détaillés du recensement peut comporter, quant à son caractère représentatif de l'ensemble des logements de la commune, de petites fluctuations liées au poids variable des adresses de grande taille et des adresses nouvelles.

1 000 à <2 000	4%
2 000 à < 3 000	3%
3 000 à < 5 000	2,5%
5 000 à < 10 000	2%
10 000 à <20 000	1,6%
20 000 ou plus	< 1%

Supposons qu'un résultat produit (nombre d'ouvriers qualifiés, par exemple) soit égal à 800. Le coefficient de variation de l'estimation est donc de l'ordre de 5%, ce qui signifie que l'écart-type vaut $800 \times 0,05$, c'est-à-dire 40. L'intervalle de confiance du résultat, habituellement construit en prenant une marge d'incertitude de plus ou moins deux écarts-type est donc : 800 ± 80 . Exprimé autrement : la grandeur cherchée a 95 chances sur 100 de se situer dans l'intervalle allant de 720 à 880.

A titre d'illustration, on peut dresser une liste d'exemple de statistiques sur les communes ayant servi à l'étude. Le tableau suivant donne, pour diverses statistiques, le résultat de 1999, la marge d'incertitude qui serait celle du sondage (arrondie à la dizaine) et la valeur des mêmes variables en 1990.

Variable	Valeur 1999	Coeff. de variation	Marge	Valeur 1990
<i>GRENOBLE</i>				
Ingénieurs, cadres	6 906	2%	± 280	5 971
Adresses achevées entre 1949 et 1974	3 003	2,5%	± 150	3 122
<i>LYON 6</i>				
Ouvriers qualifiés	714	5%	± 70	935
Divorcés	2 714	3%	± 160	2 191
<i>LYON 8</i>				
Ingénieurs, cadres	2 114	3%	± 120	1 911
<i>ROMANS</i>				
Logements d'une pièce	772	5%	± 80	821
Etrangers	2 697	3%	± 160	2 254

Ainsi, entre 1990 et 1999, le nombre d'ingénieurs, cadres à Grenoble s'est accru de près d'un millier. C'est pourtant la valeur 5 971 qui est utilisée à partir de 1990 jusqu'au recensement suivant, neuf ans plus tard, où l'on constate l'évolution. Avec la nouvelle méthode, l'estimation, en 1999, même avec une marge d'incertitude de 280, est plus en rapport avec la réalité. A contrario, l'évolution du nombre de logements d'une pièce à Romans est de sens indéterminé puisque la baisse apparente est inférieure à la marge de précision de l'instrument. Il faut garder à l'esprit, dans l'interprétation des résultats, la précision de l'instrument de mesure.

Remarque

Le détail d'information à la fois utile et fiable au niveau infracommunal mérite réflexion. Les tableaux des recensements généraux de la population offraient un grand détail de données au niveau de l'IRIS et même de l'îlot. Principalement utilisés pour la recomposition de territoires plus vastes, ils n'en conduisaient pas moins certains utilisateurs à être abusés par l'illusion de précision d'un nombre absolu et à s'en servir à des niveaux géographiques très

fins. Or, même exhaustif, le recensement n'était pas « parfait » et pouvait souffrir de certaines incertitudes, tant en ce qui concerne le dénombrement de la population que, plus encore, la déclaration précise de ses caractéristiques. De plus, la stabilité temporelle de données n'est naturellement pas assurée au fil des ans sur des groupes de petit effectif et donc très sensibles aux comportements individuels.

D.4.3 Mode d'élaboration des résultats des enquêtes annuelles de recensement

D.4.3.1 Les données et leur provenance⁷⁴

Chaque année, les enquêtes de recensement concernent un cinquième des communes de moins de 10 000 habitants et un échantillon d'adresses dans chacune des autres communes.

Pour les premières, il s'agit d'un recensement « traditionnel » où l'on recueille exhaustivement les informations de base sur la population et les logements. Ce recueil d'information porte aussi sur toutes les personnes vivant en communauté, ou dans des habitations mobiles, ou qui sont sans abri. Ces communes sont réparties de sorte que, dans chaque région, chaque cinquième soit une image fidèle de l'ensemble des communes de moins de 10 000 habitants de la région en matière de critères démographiques et de répartition des logements. Dans les départements d'outre-mer, le faible nombre de communes de moins de 10 000 habitants ne permettant pas une telle répartition, les groupes de rotation ont été constitués de façon à équilibrer la charge de collecte.

Dans les communes de 10 000 habitants ou plus, les adresses ont été préalablement réparties en cinq groupes, dits groupes de rotation, conçus pour que chacun soit une bonne représentation de sa commune. Les groupes sont sondés à tour de rôle en cinq ans de sorte que le pourcentage de logements de la commune (et donc de population) enquêtés chaque année soit égal à 8 %. Il est tenu compte, dans ces opérations, d'une stratification des adresses selon leur taille : les adresses les plus grandes comprenant plusieurs dizaines de logements (au moins 60 logements par adresse) et qui, ensemble, représentent au plus 10% des logements de la commune sont visitées exhaustivement en cinq ans. Enfin, les adresses nouvelles, c'est-à-dire construites au fil des ans, sont elles aussi toutes enquêtées puisqu'elles n'existaient pas lors du recensement général de 1999.

Rappelons aussi que les populations des communautés, dans les communes de 10 000 habitants ou plus, sont enquêtées tous les cinq ans, par roulement des communes, et que les autres populations (populations vivant habituellement dans des habitations mobiles et personnes sans abri) sont comptées une fois en cinq ans.

Le traitement des enquêtes annuelles consiste donc à rassembler toutes ces données et à les mettre en forme pour des descriptions statistiques rapides mais complètes aux niveaux national et régional. Les développements qui suivent exposent les méthodes qui permettent d'extrapoler les données des communes de moins de 10 000 habitants et celles des échantillons des autres communes.

⁷⁴ Ce qui suit est un bref rappel des dispositions décrites au [chapitre B](#).

D.4.3.2 Le traitement de la population des résidences principales des communes de moins de 10 000 habitants

Pour ces communes, le poids initial des données les concernant (c'est-à-dire l'inverse du taux de sondage) est égal à 5, puisque chaque groupe a été défini pour être un cinquième de sa région. Il faut cependant corriger légèrement ce poids pour tenir compte des évolutions. Ainsi, même si les groupes évoluent à peu près au même rythme, on ne peut pas garantir que la proportion d'un cinquième de la population est en permanence exactement respectée ; il arrive aussi que des communes fusionnent, ou se séparent, perturbant un peu les équilibrages initiaux ; il arrivera aussi que des communes franchissent, à la hausse ou à la baisse, le seuil des 10 000 habitants et changent donc de catégorie. Dans ces divers cas, on fait évoluer les groupes de rotation de façon à maintenir au mieux les équilibrages.

Comment calcule-t-on les résultats ? Chaque année, on connaît exactement la répartition annuelle des communes. On connaît aussi les résultats en nombre de résidences principales et en population établis au recensement précédent (pour la première période quinquennale, il s'agit du recensement de 1999). Selon ces résultats, on est à même de calculer précisément la proportion régionale en résidences principales représentée par le groupe de l'année et c'est cette proportion qui permet de définir le bon poids d'extrapolation.

Par exemple, si une année donnée on établit que, selon le dernier recensement⁷⁵, les communes d'un groupe comprenaient 21% des logements de toutes les communes de moins de 10 000 habitants de la région, les données du groupe sont pondérées par le coefficient $1/0,21 = 4,762$.

Remarque : pour les descriptions concernant les logements autres que principaux (résidences secondaires, logements vacants et logements occasionnels), on établit un poids spécifique. En effet, les particularités régionales font que les groupes de communes ne sont pas équilibrés selon ces catégories particulières de logements. La méthode est analogue à celle exposée plus haut : on détermine la part régionale du groupe de l'année pour cette catégorie de logements selon les résultats du dernier recensement et on en déduit le coefficient d'extrapolation.

D.4.3.3 Le traitement de la population des résidences principales des communes de 10 000 habitants ou plus

D.4.3.3.1 Les communes de métropole

Les données recueillies sont pondérées en fonction du plan de sondage appliqué dans les communes. Il convient donc de considérer séparément les adresses selon la strate à laquelle elles appartiennent. Le RIL (répertoire d'immeubles localisés) utilisé comme base de sondage pour la commune est réparti selon ces strates : adresses de grande taille, adresses nouvelles, autres adresses.

Traitement des « autres adresses »

Pour comprendre le mécanisme de pondération, prenons un exemple chiffré. Soit une commune de 25 000 logements, dont 2 000 sont répartis dans 20 adresses de grande taille

⁷⁵ Le dernier recensement doit s'entendre comme le recensement de 1999 puis, à partir de 2008, comme la somme des observations des cinq dernières enquêtes de recensement, toutes ramenées à l'année médiane de ces cinq années.

de chacune 100 logements et 1 000 répartis dans 100 adresses nouvelles de chacune 10 logements. Le RIL contient donc une strate de 20 adresses de grande taille recensées en cinq ans à raison d'un groupe de quatre adresses par an, une strate de 100 adresses nouvelles recensées en cinq ans à raison de 20 adresses par an et une strate d'autres adresses comprenant au total 22 000 logements. Cette strate est fractionnée en cinq groupes dont chacun comporte 4 400 logements (22 000 / 5).

L'échantillon de l'année porte sur 8% du total de la commune, soit 2 000 logements (25 000 x 0,08). Il est constitué de 400 logements enquêtés dans la strate des adresses de grande taille et 200 logements dans celle des adresses nouvelles et donc de 1 400 logements dans les autres adresses. Le tableau suivant résume la situation.

Strate	Nombre de logements	Groupe annuel	Echantillon annuel
Adresses de grande taille	2 000	400	400
Adresses nouvelles	1 000	200	200
Autres adresses	22 000	4 400	1 400
Total	25 000	5 000	2 000

Les unités de l'échantillon annuel dans la strate des autres adresses sont donc pondérées par le coefficient $22\ 000 / 1\ 400 = 15,714$.

A partir de 2009, les coefficients d'extrapolation seront annuellement actualisés à partir des résultats détaillés des recensements.

Traitement des adresses de grande taille

En toute logique, on devrait appliquer le même calcul pour les adresses de grande taille, mais cela n'est pas possible pour un certain nombre de communes. Les adresses de grande taille sont définies – on l'a vu - de telle sorte qu'elles comportent au moins 60 logements chacune et que leur poids total n'excède pas 10% des logements de la commune. Il s'ensuit que pour un certain nombre de communes l'application des critères aboutit à une strate contenant moins de cinq adresses de grande taille. Dans ces communes, il y aura donc des années dans lesquelles la strate ne sera pas représentée.

La solution est d'élargir le calcul à un ensemble plus vaste, à savoir la région, puisque c'est le niveau géographique de base pour la diffusion des résultats. L'exemple suivant illustre la procédure. Supposons une région avec trois communes de plus de 10 000 habitants. Deux d'entre elles, comme dans l'exemple précédent, avec 20 adresses de 100 logements réparties en cinq groupes annuels, la troisième avec 4 adresses de 80 logements chacune, réparties dans un cycle de cinq ans avec une année « vide ». Le tableau suivant résume la situation cette année-là.

Commune	Adresses de grande taille	Total des logements	Logements du groupe annuel
A	20	2 000	400
B	20	2 000	400
C	4	320	0
Total région	44	4320	800

En définitive, les 800 logements enquêtés cette année-là dans la région sont extrapolés avec le coefficient $4\,320 / 800 = 5,4$.

L'hypothèse, réaliste, qui est faite est que, même si une année donnée l'échantillon n'a pas d'adresse de grande taille dans certaines communes, il y en a suffisamment dans toute la région pour asseoir une statistique fiable.

D.4.3.3.2 Le cas particulier des communes des DOM

La méthode doit être adaptée aux communes des DOM pour lesquelles, d'une part, on ne dispose pas encore de RIL et où, d'autre part, il n'y a que très peu d'adresses de grande taille.

Le territoire d'une commune est découpé en îlots, répartis en cinq groupes annuels de sorte que chaque groupe représente une image fidèle de la commune pour les principaux critères démographiques. A tour de rôle, ces groupes font d'abord l'objet d'une enquête cartographique destinée à localiser la totalité des bâtiments et à en dénombrer précisément les logements. Dans le groupe d'une année, on réalise l'enquête de recensement de la population au sein d'un échantillon de 40 % des logements ainsi identifiés, ce qui ramène, comme en métropole, à un taux de sondage annuel de 8 % pour l'ensemble de la commune. La procédure d'extrapolation est alors la suivante : les logements enquêtés sont d'abord pondérés pour s'ajuster au nombre exact de logements du groupe annuel (coefficient de l'ordre de 2,5), puis extrapolés à la commune entière.

D.4.3.4 Fichier récapitulatif

Au total, le fichier des données recueillies auprès de la population des résidences principales rassemble le détail de tous les échantillons de l'année :

- les données individuelles provenant des communes de moins de 10 000 habitants, pondérées selon le poids du groupe de l'année dans l'ensemble ;
- les données individuelles provenant des 8% de logements enquêtés dans les autres communes, avec des poids spécifiques en métropole selon la catégorie des adresses auxquelles appartiennent ces logements (adresses de grande taille, adresses nouvelles, autres adresses). Dans les départements d'outre-mer un seul jeu de poids est utilisé, toutes les adresses entrant dans la catégorie « autres ».

Remarques :

- 1 - Les coefficients associés aux éléments des « autres adresses » sont calculés commune par commune.
- 2 - Dans les communes de 10 000 habitants ou plus, le principe des extrapolations est fondé sur le nombre total de logements. En métropole, il peut arriver que, dans le RIL, le nombre de logements de certaines adresses ne soit pas connu avec certitude. C'est essentiellement le cas des adresses nouvelles tant qu'elles n'ont pas été recensées. Pour pallier le risque de biais qui en résulterait, on effectue la correction suivante : on compare le nombre exact de logements des adresses nouvelles recensées dans l'année avec le nombre calculé à l'aide du RIL et on extrapole la différence à l'ensemble des adresses nouvelles. Par exemple, si les adresses recensées ont à elles toutes un nombre de logements inférieur de 1% à ce qui figurait dans le RIL, on considérera que ce 1% s'applique à l'ensemble des adresses nouvelles du RIL, puisque l'échantillon a été réparti aléatoirement sur tout le territoire.

D.4.3.5 Les autres populations

D.4.3.5.1 Les communautés

Pour les communes de moins de 10 000 habitants, les communautés sont recensées tous les cinq ans, dans le groupe de rotation des communes auxquelles elles appartiennent. Les résultats annuels sont extrapolés à l'ensemble des communes de moins de 10 000 habitants de la région de la même façon que la population des résidences principales, avec les mêmes coefficients.

Dans les communes de 10 000 habitants ou plus, le problème est différent. Ces communes sont bien réparties en cinq ensembles pour réaliser la collecte, mais il n'est pas possible d'équilibrer ces ensembles de sorte que chacun soit représentatif de l'ensemble des communes de 10 000 habitants ou plus pour la population vivant dans les différents types de communautés. La solution retenue en régime permanent est de maintenir inchangées pendant cinq ans (jusqu'à la prochaine collecte) les données recueillies une année dans une commune.

Par exemple, dans une région donnée dont on établira la description statistique en 2015, la population totale des communautés sera estimée à partir des données du groupe recensé en 2015, auxquelles on adjoindra celles recueillies dans les quatre autres groupes les années précédentes.

En d'autres termes, on gère un fichier exhaustif des populations des communautés que l'on actualise par cinquième tous les ans par roulement.

Remarque :

Pour les années du premier cycle, une disposition particulière des calculs sera mise en œuvre pour tenir compte du fait que, par exemple en 2004, on aura les données des communautés recueillies dans le groupe de communes de cette année et que les informations les plus récentes des autres groupes dateront de 1999 avec un questionnaire de recensement légèrement différent.

D.4.3.5.2 Les personnes résidant en habitation mobile, les personnes sans abri et les mariniers

Pour ces catégories de population, la solution la plus simple est de conserver les données collectées lors d'un recensement jusqu'au suivant, cinq années plus tard.

Dans les communes de moins de 10 000 habitants, les personnes résidant en habitation mobile et les personnes sans abri sont recensées dans leur groupe de rotation alors que, dans toutes les autres communes, un comptage simultané a lieu tous les cinq ans, à partir de l'année 2006.

Sur l'ensemble du territoire, les mariniers font eux aussi l'objet d'un comptage simultané tous les cinq ans, à partir de l'année 2006.

Remarque :

Pour toutes les communes, le comptage est initialisé avec le recensement de 1999. Comme pour les communautés, il s'agira d'un ensemble rafraîchi par parties tous les ans.

Questions complémentaires

Question 1 : *Quelle différence entre les résultats de population légale et le bilan démographique publié chaque année par l'Insee ?*

Réponse : Le bilan démographique permet, dès les premiers mois de l'année, de donner la population totale de la France et d'analyser la fécondité et la mortalité du pays. Il est élaboré à partir de sources administratives, notamment l'état civil.

Contrairement aux résultats fournis par le recensement de la population, le bilan démographique ne renseigne pas sur le niveau et la structure de la population aux échelons locaux : régions, départements, *a fortiori* communes. Seul le recensement fournit de telles informations.

Le bilan démographique est publié par l'Insee chaque année, en début d'année. Il ne faut pas confondre ce bilan avec les collectes d'informations du recensement qui sont en cours chaque année à la même époque. Les résultats issus des cinq enquêtes de recensement des années A-4 à A servent notamment à produire le chiffre définitif de la population au 1^{er} janvier de l'année A-2, base des bilans démographiques ultérieurs.

Question 2 : *La "brique de base" pour la diffusion des résultats des enquêtes annuelles de recensement est la région. Est-il possible d'obtenir des statistiques sur des territoires autres, par exemple l'ensemble des communes concernées par la loi sur le littoral, l'ensemble des zones urbaines sensibles (ZUS) de la métropole, l'ensemble du territoire rural, etc. ?*

Réponse : Les exploitations sur des zonages à la demande sont possibles, sous certaines conditions relatives à la taille et à la structure des zones concernées. Comme on l'a vu, la répartition des communes et le traitement des données sont conçus pour optimiser les descriptions régionales. Les statistiques construites sur d'autres territoires auront, par construction, des intervalles de confiance moins étroits.

Dans le cas d'une zone englobant des communes de moins de 10 000 habitants, encore faut-il qu'elles soient suffisamment nombreuses pour ne pas être trop inégalement réparties dans les groupes de rotation. Le cas des communes de 10 000 habitants ou plus est *a priori* moins contraignant puisqu'elles bénéficient toutes d'un échantillon annuel. Néanmoins, la masse d'information recueillie doit être assez grande pour avoir une bonne précision.

Les simulations effectuées sur la base du recensement de 1999 montrent qu'une taille raisonnable pour réaliser des exploitations avec une bonne précision doit être de l'ordre de celle d'une région, c'est-à-dire plusieurs centaines de milliers de personnes.

D.4.3.6 Précision des résultats des enquêtes annuelles de recensement

Comme on l'a vu, le plan de sondage des enquêtes annuelles est optimisé au niveau de la région (les groupes de rotation des communes de moins de 10 000 habitants sont définis régionalement) et les variables d'équilibrage sont principalement de nature démographique.

Les résultats des enquêtes annuelles au niveau régional fourniront donc de très bons résultats pour la pyramide des âges ou la population active par exemple. Ces résultats seront disponibles avant la fin de l'année de l'enquête et fourniront une information « avancée » très récente avec un bon niveau de précision pour la mise à jour des principales grandeurs statistiques régionales. La précision des résultats issus de l'échantillon annuel dépend en effet de la répartition par taille des communes de moins de 10 000 habitants, mais aussi de la part des communes de 10 000 habitants ou plus dans la région. A cet

égard, des abaques régionaux de précision seront publiés. De plus, les résultats de l'enquête annuelle correspondent à des données collectées à la même date, ce qui peut permettre de mesurer des effets « conjoncturels » qui seront lissés par le cumul quinquennal des résultats détaillés.

D.5 De nouvelles utilisations des résultats statistiques

D.5.1 Types d'utilisation des résultats du nouveau recensement

Les développements qui suivent mettent en évidence les types de résultats apportés par le nouveau recensement et les niveaux de précision correspondants. L'annualisation des résultats reste évidemment un point de progrès majeur du recensement. Nourris de simulations effectuées à partir de vastes ensembles de données, des exemples d'utilisation des résultats du nouveau recensement sont ici présentés pour des problématiques auxquelles les utilisateurs sont souvent confrontés ; ces exemples mettent en évidence les différences par rapport aux recensements généraux de la population.

Une de ces différences concerne les niveaux et les structures⁷⁶.

Les résultats publiés des recensements généraux de la population étaient beaucoup plus souvent des effectifs que des structures. Or, si les niveaux sont utiles pour mesurer, par exemple, l'effectif d'une population concernée par une action publique, ils ne permettent pas facilement les comparaisons entre communes ou territoires, sauf à recalculer des proportions. De plus, s'agissant désormais d'enquêtes de recensement effectuées pour partie par sondage, la précision obtenue sur une proportion est supérieure à celle obtenue sur une donnée en niveau absolu.

C'est pourquoi davantage d'informations de structure sont introduites dans les résultats statistiques produits par le nouveau recensement, sans pour autant négliger la diffusion des effectifs les plus pertinents. Les indicateurs en valeur absolue sont en effet fondamentaux lorsqu'il s'agit de mettre en œuvre ou d'orienter l'action publique.

Un travail a été engagé pour repérer de "grands indicateurs" jugés essentiels à toutes les échelles géographiques, tant communales qu'intercommunales.

Une première grille d'analyse naturelle consiste à **distinguer les utilisations selon le niveau géographique**, en partant du niveau régional pour atteindre le niveau intercommunal. Une attention plus particulière doit être portée à certaines variables dont l'utilisation peut paraître complexe ou risquée aux utilisateurs dans le cadre du nouveau recensement. Ainsi en est-il de l'âge, en raison du caractère pluriannuel de la collecte, mais aussi des informations bilocalisées, à cause de la non-simultanéité du recueil de l'information.

Par ailleurs, l'utilisation des résultats portant sur les logements, en liaison avec le traitement particulier des **constructions nouvelles**, a fait l'objet d'un approfondissement.

⁷⁶ Voir aussi l'[encadré](#).

Encadré 11 : A propos de l'utilisation des résultats statistiques

1 - Comparaisons spatiales

L'année de collecte est utile pour effectuer des comparaisons entre les communes de moins de 10 000 habitants. En effet, afin de pouvoir les comparer, les résultats des communes de moins de 10 000 habitants seront calés sur le niveau de population de la commune estimée en A-2. Pour la majorité des critères, cette procédure sera suffisante.

En revanche, pour des variables à caractère conjoncturel comme l'emploi ou le chômage, comparer des résultats issus de données recueillies à des dates différentes peut être hasardeux. Dans ce cas, il convient de passer par l'intermédiaire d'une référence régionale ou nationale. Par exemple, le taux de chômage de chaque commune peut être rapporté au taux de chômage régional de l'année de collecte, puis cet indicateur de sur- ou sous-exposition au chômage peut être comparé entre les communes.

Les communes de 10 000 habitants ou plus ne posent pas ce problème car elles sont enquêtées tous les ans dans les mêmes conditions ; leurs résultats sont donc strictement comparables entre eux.

2 - Comparaisons temporelles

La comparaison de résultats publiés chaque année peut conduire l'utilisateur à comparer abusivement une année à la précédente. S'agissant des résultats des enquêtes annuelles, il faut s'assurer que l'incertitude liée au sondage n'est pas supérieure à l'ampleur de l'évolution. Pour les résultats détaillés issus des cinq années, le rafraîchissement d'une année sur l'autre ne porte que sur un cinquième des communes de moins de 10 000 habitants et sur un cinquième de l'échantillon dans les communes de 10 000 habitants ou plus. Il est donc plus judicieux d'effectuer les comparaisons à cinq ans d'intervalle, sur un ensemble complètement renouvelé.

3 - Mise en cohérence avec d'autres sources

Le nouveau recensement, pas plus que l'ancien, ne pourra répondre à l'ensemble des besoins d'information exprimés. Le recensement général fournissait tous les huit à neuf ans une base sur laquelle des estimations étaient élaborées à partir d'autres sources, le plus souvent administratives, au cours de la période intercensitaire. Avec le nouveau recensement, ce dispositif est appelé à changer sensiblement. Le recensement quitte son rôle de référence décennale immobile pour offrir une actualisation régulière. Bien adapté au suivi des phénomènes à évolution relativement lente, comme le sont la plupart des phénomènes sociodémographiques, il ne peut évidemment se substituer au suivi conjoncturel que permettent des sources telles que les statistiques de demandes d'emploi de l'ANPE pour la mesure du chômage ou les DADS pour celle de l'emploi salarié. Il faut donc construire un nouveau paysage statistique local en donnant à chaque source sa place en cohérence avec les autres.

D.5.1.1 Utilisations selon le niveau géographique

D.5.1.1.1 Utilisations aux niveaux national et régional

Les résultats du recensement (résultats détaillés issus des cinq enquêtes de recensement les plus récentes) seront utilisables sans précaution particulière, si ce n'est de bien mesurer qu'ils sont représentatifs, en moyenne, de l'année médiane de ces cinq dernières années. Publiés en fin d'année N, ils porteront donc sur le début de l'année N-2. Ils seront directement comparables aux résultats que produisait un recensement général.

Toutefois, rappelons que le plan de sondage des enquêtes annuelles est optimisé au niveau de la région (les groupes de rotation des communes de moins de 10 000 habitants sont définis régionalement) et que les variables d'équilibrage sont principalement de nature démographique.

D.5.1.1.2 Utilisations sur des territoires multicommunaux

Les résultats détaillés portent sur les cinq dernières enquêtes de recensement. Pour un territoire multicommunal, toutes les communes de moins de 10 000 habitants qui le composent ont été recensées au cours des cinq dernières enquêtes, et on dispose d'un échantillon de 40% de la population et des logements de ses communes de 10 000 habitants ou plus. Les données ainsi recueillies sont donc représentatives en moyenne du milieu de la période, à savoir l'année N-2. Pour la majorité des comportements démographiques, assez inertes, le décalage temporel de collecte selon les communes entre les cinq années n'apporte pas de perturbation significative. De plus, par comparaison avec les résultats d'un recensement général, utilisés de nombreuses années après sa collecte, l'utilisateur n'est pas conduit à des conclusions erronées. En revanche, le caractère annuel de la mise à jour des données peut légitimement inciter à examiner les résultats pour des variables nettement plus conjoncturelles comme l'emploi ou le chômage. Dans ce cas, le décalage temporel, même mineur au sein d'une période de cinq années, peut induire en erreur l'utilisateur qui n'aurait pas pris quelques précautions.

On montre cependant que l'on peut en déduire une vision, certes lissée, de l'évolution de l'emploi dans une zone d'emploi, et ainsi suivre par exemple l'impact d'une restructuration industrielle. De plus, en tenant compte de l'année de recensement des différentes communes, il est possible de réaliser entre communes des comparaisons spatiales de niveaux relatifs d'emploi et de chômage. Pour ce faire, il suffit de comparer la commune à sa référence régionale de l'année, connue au moyen de l'enquête annuelle, en calculant par exemple un indice de sur- ou sous-exposition au chômage. Les indices ainsi calculés sur la période des cinq dernières années peuvent être comparés entre eux en s'affranchissant ainsi des différences conjoncturelles de situation de l'emploi au cours de la période. Dans tous les cas, l'information est nettement plus riche qu'avec un recensement général, qui ne produisait une photographie que tous les huit ou neuf ans, sans considération de cycle conjoncturel.

D.5.1.1.3 Utilisation aux niveaux communal et infracommunal

C'est au niveau communal que la différence de traitement entre les communes au-dessous ou au-dessus du seuil de 10 000 habitants a le plus d'implications. En effet, les communes de moins de 10 000 habitants, recensées exhaustivement tous les cinq ans, connaîtront une situation assez proche de celle qui prévalait lors des recensements généraux. L'intervalle intercensitaire étant réduit à cinq ans, elles bénéficient cependant d'une actualisation plus fréquente, mais ponctuelle. En revanche, dans les communes de 10 000 habitants ou plus, la collecte et donc le rafraîchissement sont annuels. Toutefois, les résultats qui en découlent

sont d'une nature différente de ceux, ponctuels, des recensements généraux et nécessitent donc une certaine accoutumance pour leur faire donner leur pleine mesure.

En particulier, le recours à l'échantillonnage, même à un taux élevé de 40% sur cinq ans, a marqué une rupture majeure dans l'esprit des utilisateurs. Pourtant, il faut rappeler que les recensements généraux n'étaient jamais exploités exhaustivement pour toutes les variables et que certaines variables, comme la profession et la catégorie socioprofessionnelle, mais aussi l'activité économique ou les structures familiales, n'étaient exploitées que pour un quart des individus. Dans le nouveau recensement, les variables profession, activité économique et structures familiales sont exploitées, comme par le passé, sur un quart des questionnaires des communes de moins de 10 000 habitants, mais sur la totalité des questionnaires collectés dans les communes de 10 000 habitants ou plus. Pour ces communes, on atteindra donc un taux de 40% pour ces informations au lieu de 25% par le passé.

En second lieu, les données du recensement général, portant toutes sur une même date, présentaient un caractère plus concret et rassurant que l'accumulation de données pendant cinq ans. Il incombe donc aux statisticiens-producteurs de prouver aux utilisateurs l'utilité, pour leurs besoins habituels, de résultats construits à partir de ces données collectées au cours de cinq années et notamment de montrer que, par rapport à des résultats ponctuels mais utilisés plusieurs années après, des informations pluriannuelles présentant un caractère lissé permettent de mieux saisir les tendances de fond au delà des bruits conjoncturels et de suivre les mouvements en cours plutôt que de les constater *ex post*.

Mesure des mouvements migratoires

La connaissance des mouvements migratoires au niveau communal est une préoccupation majeure des communes. Le recensement général ne permettait cette mesure que tous les huit ou neuf ans, donc à des intervalles variables, et induisait une analyse fortement normative de ces mouvements, en les considérant stables sur la période et en occultant par exemple les phénomènes de retournement.

Le nouveau recensement mesure les migrations de la même façon qu'auparavant : on calcule le solde migratoire dit « apparent » en partant de l'évolution de la population, dont on retranche l'accroissement naturel (les naissances moins les décès) tel que le fournissent les statistiques de l'état civil. Le fait de pouvoir réaliser ce calcul à intervalles rapprochés permet de repérer plus rapidement les inflexions éventuelles des tendances. Compte tenu du caractère cyclique du nouveau recensement, il est conseillé de calculer ces soldes sur un intervalle de cinq ans, mais on peut actualiser ce calcul tous les ans.

Pour intégrer la non-simultanéité de la collecte des données, on fait l'hypothèse supplémentaire que les probabilités de migration ont une certaine inertie dont les évolutions ne sont significativement décelables que sur des périodes d'une certaine longueur. La conséquence de cette hypothèse est que, comme pour les autres variables du recensement, les résultats produits donnent une vision « moyenne » de la période et que les évolutions sont lissées. Cette façon de modéliser est usuelle en démographie, par exemple dans l'analyse de la fécondité.

Typologie d'IRIS

Une des utilisations des résultats du recensement au sein d'une commune est la cartographie sociale de ses quartiers. Il s'agit le plus souvent d'établir une typologie des différents quartiers au regard de critères comme la structure sociale de la population, l'exposition au chômage, les caractéristiques du parc de logements, toutes caractéristiques qui présentent une assez grande inertie. Il est donc légitime de s'interroger sur la robustesse d'une telle typologie à l'échantillonnage dans les villes comptant au moins 10 000 habitants.

En effet, quel serait l'intérêt d'une actualisation annuelle si l'aléa engendré par le sondage masquait les évolutions réelles des quartiers ?

Des simulations réalisées sur des IRIS du recensement de 1999 montrent que le classement des IRIS n'est pas affecté par le recours au sondage. De plus, cette robustesse n'est pas affectée, bien au contraire, par la complexité des critères pris en compte pour réaliser la typologie. En revanche, ce type d'analyse peut désormais être mené à tout instant, sur la période des cinq dernières années, sans considération de la date du dernier recensement.

Zones à la demande

Le nouveau recensement, parce qu'il s'appuie non plus sur la brique de base que constituaient les îlots en ville, mais sur le répertoire des adresses, offre la possibilité de recomposer des territoires infra-urbains, non plus en termes d'îlots, mais en termes d'adresses. Ainsi devient-il possible de produire des informations nouvelles, comme par exemple, dans le cas d'un projet de tramway, le nombre de personnes résidant le long d'un tracé envisagé.

Un autre exemple est celui de la directive européenne sur le bruit, qui fait obligation aux pays membres d'établir des cartes d'exposition au bruit de la population dans les agglomérations de plus de 250 000 habitants à partir de 2007, puis de plus de 100 000 habitants à partir de 2012. A partir de la caractérisation des voies en termes d'exposition au bruit, il est possible d'estimer avec une précision tout à fait acceptable la population résidente le long des axes concernés.

Autre exemple encore, l'analyse des caractéristiques de la zone de chalandise d'un projet d'implantation commerciale est nécessaire pour l'examen du projet en commission départementale d'équipement commercial (CDEC). Dans le cas d'un équipement urbain, la zone de chalandise est en général limitée à la proximité immédiate du commerce. Faute d'avoir des résultats sur cette zone spécifique, il peut être utile de disposer au minimum d'une clé de répartition des résultats des IRIS englobants, par exemple du nombre de logements ou de personnes au niveau très fin de l'îlot ou même de l'adresse.

Cas des données bilocalisées

Le cas des données de flux est sans doute celui qui interpelle le plus les utilisateurs. Comment obtenir la même information qu'au recensement général où les flux, par exemple de déplacements de commune à commune pour raison de travail, étaient saisis au même instant ? Les communes étant recensées à des dates différentes, les déplacements d'une commune A vers une commune B seront observés une année donnée, alors que les flux réciproques, à savoir de la commune B vers la commune A seront peut-être observés quatre ans plus tard !

Une simulation a été réalisée sur l'agglomération de Montpellier. La typologie des déplacements domicile-travail au sein de cette agglomération montre une très forte polarisation autour de la ville-centre. Par ailleurs, dans une agglomération de cette importance, la répartition des communes de moins de 10 000 habitants entre les groupes de rotation de communes présente un caractère tout à fait équilibré. Ainsi, tous les ans, 45 000 habitants sont recensés en moyenne dans l'agglomération, actualisant régulièrement les informations.

Pour fixer les idées sur la précision, considérons à titre d'exemple le nombre d'habitants de Lattes qui se déplacent pour aller travailler à Montpellier. Il était de l'ordre de 2 700 personnes en 1999. Si ce niveau reste inchangé, l'incertitude sera de l'ordre de 3,5% (en termes de coefficient de variation). Ce qui signifie qu'une évolution annuelle moyenne de 1,5% du nombre de personnes allant travailler à Montpellier sera jugée significative au bout

de 5 ans⁷⁷. L'actualisation des informations relatives aux déplacements domicile-travail est plus rapide que les 8 ou 9 ans qu'il y avait entre les anciens recensements généraux. Le recensement général apportait certes la connaissance « anecdotique » des tout petits flux, mais ce sont précisément ces flux qu'il fallait souvent filtrer dans les tableaux afin de ne pas en compliquer la lecture !

D.5.1.2 Les constructions nouvelles

Dans les communes de 10 000 habitants ou plus, le plan de sondage tient compte spécifiquement des adresses nouvelles, dont il est important de connaître le nombre exact de logements afin de réaliser les extrapolations. A cet effet, les adresses nouvelles sont enquêtées exhaustivement dans les cinq ans au maximum qui suivent leur entrée dans le RIL.

Chaque année, l'Insee dispose sur les adresses nouvelles de données exhaustives datées de trois ans en moyenne. Dès lors, une caractérisation statistique fine et annuelle des nouveaux habitants est envisageable au bénéfice des utilisateurs. C'est un apport non négligeable du nouveau recensement pour l'étude de la construction neuve, qui reste le plus souvent un phénomène conjoncturel à l'échelle d'une commune (un programme immobilier concentre généralement la construction neuve sur quelques années). Les recensements généraux, qui avaient lieu tous les 8 ou 9 ans, ne permettaient pas d'appréhender de manière totalement satisfaisante la construction neuve.

De plus, une des composantes nécessaires aux projections de population scolaire est l'estimation du nombre d'enfants à scolariser par niveau dans les constructions neuves. Les résultats du nouveau recensement apporteront une information fine pour l'estimation des taux de peuplement et de scolarisation.

D.5.2 Synthèse

Les deux types de résultats statistiques obtenus grâce à la rénovation du recensement sont issus, les uns de cinq années d'enquête (ce sont les résultats du recensement) et les autres d'une seule enquête annuelle (ce sont les résultats de l'enquête annuelle de recensement). On l'a vu, chaque année A, l'enquête de recensement permet d'obtenir des résultats avancés pour la France et chacune de ses régions, et des résultats détaillés pour un cinquième des communes de moins de 10 000 habitants (celles enquêtées exhaustivement dans l'année). Les résultats du recensement sont, eux, des résultats détaillés produits l'année A à tous les niveaux géographiques, jusqu'à l'IRIS en infracommunal, et ils traduisent la situation moyenne des cinq dernières années A-4 à A, vue de l'année médiane A-2.

Le tableau ci-après précise les résultats à utiliser et les précautions d'emploi en fonction du niveau géographique choisi.

⁷⁷ Cinq fois 1,5%, cela donne au bout de cinq ans une croissance qui, exprimée en pourcentage, excède largement les bornes de l'intervalle de confiance de l'estimation, placées à $2 \times 3,5 = 7\%$ de la valeur de celle-ci.

**Tableau de synthèse des types de résultats statistiques
en fonction du niveau territorial**

Niveau géographique	Résultats produits fin année N	Date de référence
France, région	Résultats de l'enquête annuelle de recensement	Début de l'année N
	Résultats détaillés issus des cinq dernières enquêtes de recensement	Début de l'année N-2
Commune de moins de 10 000 habitants	Résultats détaillés de l'enquête de recensement	Début de l'année de l'enquête
	Résultats détaillés : si besoin est, actualisation annuelle des données de la dernière enquête de recensement pour les caler sur le niveau de population estimé en N-2	Début de l'année N-2
Commune de 10 000 habitants ou plus	Résultats détaillés issus des cinq dernières enquêtes	Début de l'année N-2
Ensemble multicommunal	Résultats détaillés issus des cinq dernières enquêtes : - résultats détaillés N-2 pour les communes de 10 000 hab. ou plus de la zone ; - résultats ramenés à N-2 pour les communes de moins de 10 000 hab. de la zone	Début de l'année N-2

**Propositions d'utilisation des résultats statistiques portant sur l'année N
en fonction du niveau territorial et du thème**

Niveau géographique – thèmes	Source	Précautions	Disponibilité
Région			
Population	Résultats de l'enquête annuelle de recensement	A comparer aux autres régions et à la France seulement	Fin de l'année N
	Résultats détaillés du recensement	Aucune	Fin de l'année N+2
Structures par âge	Résultats de l'enquête annuelle	Groupes d'âges	Fin de l'année N
	Résultats détaillés du recensement	Age détaillé	Fin de l'année N+2

Niveau géographique – thèmes	Source	Précautions	Disponibilité
Analyse des migrations interrégionales	Résultats de l'enquête annuelle	Se limiter aux principales régions d'échanges	Fin de l'année N
	Résultats détaillés du recensement	Aucune	Fin de l'année N+2
Population active - chômage	Résultats de l'enquête annuelle	Groupes d'âges agrégés	Fin de l'année N
	Résultats détaillés du recensement	Aucune	Fin de l'année N+2
Profession - Catégorie socioprofessionnelle	Résultats de l'enquête annuelle	Se limiter à la CS	Fin de l'année N
	Résultats détaillés du recensement	Profession détaillée	Fin de l'année N+2
Zone multicommunale (département, arrondissement, canton...)			
Tous thèmes – derniers résultats	Résultats détaillés du recensement	Aucune	Fin de l'année N+2
Mesure des évolutions : données non conjoncturelles	Résultats détaillés du recensement	Privilégier l'évolution sur 5 ans (renouvellement complet de l'information)	Fin de l'année N+2
Mesure des évolutions : données conjoncturelles : ex : suivi de l'emploi dans une zone d'emploi	Résultats détaillés du recensement	Pour un suivi annuel : vérifier les dates de recensement des communes de moins de 10 000 hab.	Fin de l'année N+2
Comparaisons entre communes de la zone : données non conjoncturelles	Résultats détaillés du recensement	Vérifier le caractère non conjoncturel	Fin de l'année N+2
Comparaisons entre communes de la zone : données conjoncturelles : ex : taux de chômage des communes de la zone	Résultats détaillés pour les communes de 10 000 hab. ou plus, dernière enquête de recensement pour celles de moins de 10 000 hab.	Calculer le rapport du taux de chômage de chaque commune de moins de 10 000 hab. à celui de la région la même année. Comparer les résultats obtenus entre les communes.	Fin de l'année N+2
Commune de moins de 10 000 habitants			
Tous thèmes – derniers résultats (y compris IRIS)	Résultats de la dernière enquête de recensement	Aucune	Fin de l'année de l'enquête
Mesure des évolutions : tous thèmes (y compris IRIS)	Résultats de la dernière enquête de recensement	Evolution sur 5 ans	Fin de l'année de l'enquête
Commune de 10 000 habitants ou plus (en régime permanent, à partir de fin 2008)			
Tous thèmes – derniers résultats	Résultats détaillés du recensement	Aucune	Fin de l'année N+2

Niveau géographique – thèmes	Source	Précautions	Disponibilité
Logements nouveaux et population correspondante	Résultats détaillés du recensement	Aucune (exhaustif)	Fin de l'année N+2
Mesure des évolutions : données non conjoncturelles	Résultats détaillés du recensement	Privilégier l'évolution sur 5 ans (renouvellement complet de l'information)	Fin de l'année N+2
Mesure des évolutions : données conjoncturelles : ex : suivi de l'emploi dans une commune	Résultats détaillés du recensement	Mesure des évolutions annuelles délicate : seul 1/5 étant rafraîchi, privilégier les évolutions quinquennales plus robustes (renouvellement complet de l'information).	Fin de l'année N+2
Comparaisons entre communes de 10 000 hab. ou plus	Résultats détaillés du recensement	Aucune	Fin de l'année N+2
Comparaisons avec une commune de moins de 10 000 hab.	Résultats détaillés du recensement	Si la variable est de nature conjoncturelle, vérifier la date d'enquête de la commune de moins de 10 000 hab.	Fin de l'année N+2
IRIS-2000 (communes de 10 000 habitants ou plus) (en régime permanent, à partir de fin 2008)			
Tous thèmes – derniers résultats – effectifs absolus	Résultats détaillés du recensement	Vérifier la précision en fonction notamment de l'effectif	Fin de l'année N+2
Tous thèmes – derniers résultats – proportions	Résultats détaillés du recensement	Plus robuste que les effectifs. Vérifier cependant la précision	Fin de l'année N+2
Logements neufs et population correspondante	Résultats détaillés du recensement	Aucune (exhaustif)	Fin de l'année N+2
Mesure des évolutions : données non conjoncturelles	Résultats détaillés du recensement	Evolution sur 5 ans (renouvellement complet de l'information)	Fin de l'année N+2
Mesure des évolutions : données conjoncturelles : exemple : suivi de l'emploi annuel	Résultats détaillés du recensement	Les résultats détaillés ne permettent pas un suivi annuel en effectif absolu mais, en revanche, ils offrent un suivi lissé en moyenne mobile quinquennale	Fin de l'année N+2
Comparaisons entre IRIS de la commune ou	Résultats détaillés du recensement	Vérifier la précision en fonction notamment	Fin de l'année N+2

Niveau géographique – thèmes	Source	Précautions	Disponibilité
d'autres communes.		de l'effectif. La précision est meilleure pour les proportions (par exemple : typologies, comparer des taux de moins de 20 ans...)	
Comparaisons entre l'IRIS et sa commune.	Résultats détaillés du recensement	Vérifier la précision en fonction notamment de l'effectif. La précision est meilleure pour les proportions (par exemple comparer le taux de moins de 20 ans...)	Fin de l'année N+2

D.6 2004-2008 : période de montée en régime de croisière

D.6.1 Particularités de la période

Comme on l'a vu, la palette des produits du recensement est riche et chacun de ces produits obéit à sa propre logique de production, de diffusion et d'utilisation :

- pour des raisons de nature juridique tenant à la nécessaire égalité de traitement des communes, les populations légales annuelles seront publiées au Journal officiel de la République française à partir de fin 2008, c'est-à-dire à partir du moment où tout le territoire national aura pu être pris en compte par les enquêtes de recensement ;
- les résultats de niveau national et régional à partir du cumul des enquêtes de recensement peuvent être produits et publiés dans les douze à quinze mois de la réalisation de cette enquête ;
- s'agissant des communes de moins de 10 000 habitants, les résultats détaillés du recensement peuvent être produits et publiés dans les quinze à dix-huit mois de la réalisation de l'enquête quinquennale et exhaustive de recensement dans ces communes ;
- en ce qui concerne les communes de 10 000 habitants ou plus, les résultats détaillés du recensement nécessitent, pour être produits et publiés chaque année, qu'on dispose de cinq enquêtes de recensement successives ; ils seront donc disponibles à partir de fin 2008 puisque la première enquête de recensement date de 2004 ; d'ici 2008, l'Insee publiera pour ces communes des estimations de population dès lors que le cumul des échantillons d'une, deux, trois ou quatre enquêtes le permettra ;
- si on considère une zone multicommunale, la diffusion des résultats du recensement entre en régime de croisière à une date qui est fonction de sa composition : fin 2008 si elle compte au moins une commune de 10 000 habitants ou plus ; possiblement à une date antérieure si elle ne regroupe que des communes de moins de 10 000 habitants et si toutes sont recensées avant cette date.

On voit clairement que le système produira chaque année une information abondante à partir de 2009 mais que, d'ici là, le nouveau recensement peut apparaître comme une machine absorbant énormément d'informations individuelles et ne restituant que peu de statistiques. D'où le choix de commencer à publier tôt, en augmentant régulièrement le volume d'informations produites et en descendant progressivement dans les niveaux géographiques.

D.6.2 Dispositions adoptées

D.6.2.1 Publier avant la fin de l'année une estimation de la population des communes recensées

Le nombre de personnes recensées lors de la collecte est communiqué à la commune à la fin de la phase de réception, d'enregistrement et de contrôle. Ce nombre correspond pour les communes de moins de 10 000 habitants à l'effectif de leur population, alors que pour les communes de 10 000 habitants ou plus il représente la population de l'échantillon enquêté.

Pour les communes, le résultat le plus attendu du recensement est bien l'effectif de la population. Même si ce chiffre ne sera authentifié qu'en 2008, il peut être diffusé sous forme provisoire comme cela se faisait lors des recensements généraux. Pour les communes de moins de 10 000 habitants, rien n'interdisait de reproduire ce schéma. En revanche, l'exercice était plus délicat pour les communes de 10 000 habitants ou plus puisqu'il fallait estimer la population des ménages à partir d'un échantillon de 8% la première année, puis 16% la deuxième, et ainsi de suite. Il a donc été nécessaire de définir un seuil minimal de population combiné au degré d'homogénéité de l'habitat permettant de produire ces estimations, seuil qui a été abaissé au fil des années au rythme de la disponibilité des données des enquêtes de recensement successives.

Le calcul des chiffres de population pour les communes concernées intervient à la fin de l'année. La diffusion se fait par mise en ligne sur le site internet <http://www.insee.fr/>. Avant la mise en ligne, une communication aux communes est réalisée, surtout pour celles de 10 000 habitants ou plus, car l'estimation ne se déduit pas simplement des comptages de la collecte.

D.6.2.2 Produire avant le démarrage de l'enquête suivante les premiers indicateurs statistiques nationaux et régionaux issus de l'enquête de recensement

A partir des premiers dénombrements sur la population et les logements, il est possible de publier au plus tard début janvier, c'est-à-dire avant le lancement de l'enquête de recensement suivante, quelques données de cadrage issues de l'enquête de recensement la plus récente.

D.6.2.3 Produire avant l'été les résultats statistiques nationaux et régionaux issus de l'enquête de recensement de l'année précédente

Dès 2007, la codification de la profession et le redressement de l'ensemble des variables des enquêtes de recensement étant devenus possibles, on produit l'ensemble des tableaux issus de chaque enquête de recensement aux niveaux France et régions. Le degré de détail et le calendrier de cette production sont agencés de manière à rejoindre progressivement le régime de croisière qui sera instauré à partir de début 2009.

D.6.2.4 Restituer aux communes avant l'été des statistiques issues de l'enquête de recensement de l'année précédente

Après les différentes phases de validation des données traitées, la fabrication de produits de restitution d'information statistique aux communes est effectuée, sous la forme de fiches présentant de manière synthétique les principaux résultats pour la commune, avec l'indication des principales évolutions depuis le recensement général de 1999 et une mise en évidence graphique.

Remarque

Pour la production de cette information au cours des premières années, les communes de moins de 10 000 habitants ne posaient pas de problème de disponibilité d'une quantité suffisante de données puisqu'elles sont recensées exhaustivement. Les caractéristiques sociodémographiques des communes recensées pouvaient en théorie être diffusées sous le format final dès la première enquête de recensement. Mais les traitements de codification et

de redressement ont fait l'objet d'une montée en charge progressive pour atteindre leur régime courant en 2006. Il a donc été nécessaire de prévoir, pour ces communes aussi, un enrichissement progressif des résultats.

En revanche, du fait du sondage, il était difficile de prévoir un dispositif uniforme pour les communes de 10 000 habitants ou plus aux cours des premières années, d'autant que seule la population des ménages était enquêtée jusqu'en 2006.

Pour les plus peuplées et les plus homogènes en terme d'habitat d'entre elles, il a été possible de produire tôt quelques résultats synthétiques comme la répartition en grands groupes d'âges ou quelques caractéristiques des logements. Pour les communes proches du seuil de 10 000 habitants, il fallait attendre d'avoir cumulé plusieurs années d'information. En effet, le poids des adresses de grande taille dans l'échantillon annuel peut perturber la significativité des résultats en introduisant des effets de grappe.

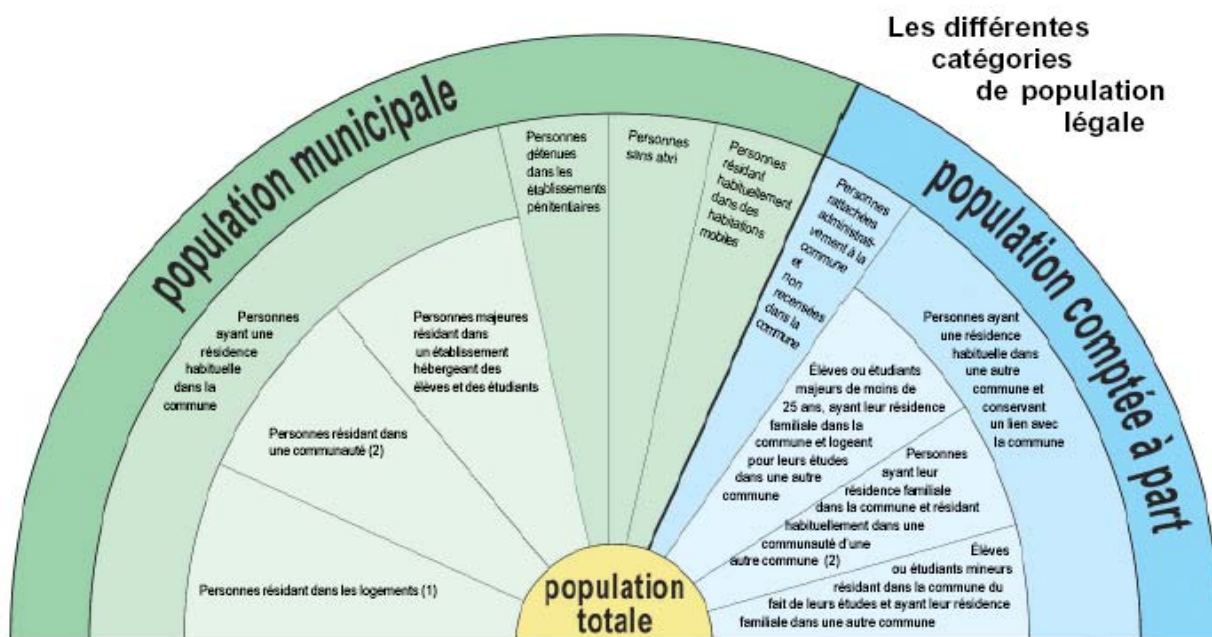
Le calendrier de restitution a donc été différencié selon la taille et l'homogénéité de l'habitat des communes. Le principe est que, progressivement, l'information est enrichie dans son contenu, la taille des communes de 10 000 habitants ou plus concernées abaissée et les délais de production contractés, de manière à atteindre fin 2008 le régime courant de production des résultats statistiques détaillés.

L'enrichissement progressif de l'information pour les communes de 10 000 habitants ou plus aux cours des premières années tient, évidemment, au fait qu'une enquête de recensement apporte chaque année de nouvelles données. Il est à noter, à ce propos, que les restitutions issues de la seule enquête de recensement de 2004 rendent compte de la situation au 1er janvier 2004 ; celles issues des deux enquêtes de recensement de 2004 et 2005 rendent compte de la situation au 1er juillet 2004 (point médian des deux enquêtes de recensement) ; celles issues des trois enquêtes de recensement de 2004, 2005 et 2006 rendent compte de la situation au 1er janvier 2005 (point médian des trois enquêtes de recensement) ; et ainsi de suite jusqu'à rejoindre, par sauts semestriels, la date du 1er janvier 2006 qui est la référence des premières productions annuelles détaillées en régime de croisière.

Après quoi, le nombre d'enquêtes de recensement utilisées demeurant égal à cinq, la date de référence se décale d'un an tous les ans. Le tableau suivant montre quelles enquêtes de recensement seront utilisées pour fabriquer les produits de diffusion des années 2004 à 2009 :

Année de diffusion	2004		2005		2006		2007		2008		2009 (régime de croisière)
	2e semestre	1er semestre	2e semestre	1er semestre	2e semestre	1er semestre	2e semestre	1er semestre	2e semestre	2e semestre	
Millésime des enquêtes de recensement utilisées	2004	2004	2004+2005		2004+2005 +2006		2004+2005 +2006+2007		2004+ 2005+ 2006+ 2007+ 2008	2005+2006 +2007+2008 +2009	

ANNEXE D 1 Catégories de population légale



(1) La résidence habituelle pour les conjoints éloignés et les élèves et étudiants mineurs, éloignés pour des raisons d'études, est la résidence familiale.

(2) À l'exception des établissements hébergeant des élèves et des étudiants (y compris les établissements d'enseignement militaire) et établissements pénitentiaires.

ANNEXE D 2

Précision des résultats du recensement au niveau infracommunal

I - Un premier indicateur de précision au niveau de l'IRIS

Le tableau présenté au chapitre D, [§ D.4.2.4](#), doit être affiné dans ses premières tranches dès lors que l'on s'intéresse au niveau infracommunal, tel que l'IRIS.

Des simulations qui respectent à la fois le plan de sondage et les principes des estimations ont permis de mettre en évidence la loi approchée suivante :

$$CV = \frac{180}{\sqrt{\text{Effectif}}}$$

qui donne le coefficient de variation en fonction de l'effectif de la variable considérée pour un IRIS. On peut ainsi calculer l'indicateur pour chaque niveau de la variable :

Effectif ou total de la variable	CV (%)
3 500	3
3 000	3,3
2 500	3,6
2 000	4
1 750	4,3
1 500	4,6
1 250	5,1
1 000	5,7
900	6
800	6,4
700	6,8
600	7,3
500	8,5
400	9
300	10,4
200	12,8
150	14,7
100	18
80	20,1
50	25,5
20	40,2
10	56,9

Exemple : l'IRIS 701 de ma commune compte 200 chômeurs. Dans ce cas, le coefficient de variation pour le nombre de chômeurs est de 12,8% pour cet IRIS.

Il convient d'avoir à l'esprit que le plan de sondage des enquêtes de recensement a été optimisé au niveau de la commune. En d'autres termes, si le total estimé de la variable d'intérêt X pour la commune est de 200 (par exemple, le nombre de chômeurs ayant entre 20 et 25 ans) et que l'on s'intéresse à une autre variable Y concernant un IRIS dont le total estimé est aussi de 200 (par exemple le nombre total de chômeurs de cet IRIS), alors la précision du total X est a priori meilleure que celle de l'estimation du total Y.

II - Comment interpréter cet indicateur ?

Quelle utilisation concrète peut-on faire de cet indicateur de précision ?

L'avantage mais aussi l'inconvénient du coefficient de variation est qu'il n'a pas d'unité ou de dimension. Autrement dit, que veut dire un coefficient de variation de 10% en général ? Quel est le moment où on estime prendre des risques ? Il n'y a pas de règle générale car cela dépend du contexte d'utilisation.

Prenons l'exemple d'une enquête par sondage, l'American Community Survey, réalisée chaque année sur l'ensemble du territoire des Etats-Unis dans l'optique de donner des caractéristiques sur la population et les logements. Le tableau qui suit concerne les résultats de l'enquête 2002 sur le comté de Santa Barbara en Californie¹²⁷.

	Effectif	CV
Hommes	193 149	0,8 %
Femmes	193 695	0,8 %
Moins de 5 ans	25 950	2,3 %
De 5 à 9 ans	23 940	12,2 %
De 10 à 14 ans	32 381	9,4 %
De 15 à 19 ans	31 044	9,8 %
De 20 à 24 ans	34 606	8,6 %
De 25 à 34 ans	51 476	3,4 %
De 35 à 44 ans	58 472	2,6 %
De 45 à 54 ans	46 950	3 %
De 55 à 59 ans	16 578	10,6 %
De 60 à 64 ans	15 102	11,9 %
De 65 à 74 ans	26 388	5,5 %
De 75 à 84 ans	16 126	9 %
85 ans ou plus	7 831	14,4 %

On voit que les résultats de l'enquête sur une population de presque 400 000 habitants peuvent comporter, pour certaines tranches d'âge, des coefficients de variation d'au moins 10%. Le contexte n'est bien entendu pas le même que celui de l'enquête de recensement en France (environ 3 millions de ménages américains sur 107 millions sont enquêtés annuellement), mais il est intéressant de se rappeler que des coefficients de variation de 10% ne sont pas rédhitoires à l'analyse et à la diffusion de résultats.

III - Préconisations

Après ces considérations générales, il convient de donner à l'utilisateur les indications lui permettant de savoir, d'une part, sur quoi il peut compter comme tableau au niveau de l'IRIS et, d'autre part, quelles précautions il doit prendre. Concrètement, il s'agit de lui proposer des règles lui permettant d'apprécier la qualité des estimations.

¹²⁷ <http://www.census.gov/acs/www/Products/Profiles/Single/2002/ACS/Tabular/050/05000US060831.htm>

La 1ère règle vient de l'effectif du croisement de variables qui intéresse l'utilisateur. Le tableau de décision qui suit lui est recommandé :

Effectif	Règle
Au-dessous de 50	Danger
Entre 50 et 199	Vigilance
Entre 200 et 500	Un peu de prudence
Au-dessus de 500	Normalement, pas de problème

Cette première règle peut être modulée par certains éléments d'appréciation. Par exemple, si les variables concernées sont fortement corrélées à l'âge, la règle est probablement trop pessimiste. En effet, les groupes de rotation d'adresses et les échantillons annuels sont équilibrés, entre autres, sur le sexe et l'âge. Cet équilibrage « profite », en termes de gain de précision, non seulement aux variables d'équilibrage elles-mêmes, mais aussi aux variables qui leur sont corrélées.

Par ailleurs, il faudra être plus vigilant sur les résultats concernant les IRIS qui ont peu d'adresses mais qui n'ont pas pour autant d'adresses de grande taille recensées exhaustivement. En effet, dans ce cas, le grain échantillonné (l'adresse en l'occurrence) est moins fin et donc la précision des estimations moins bonne.

Une 2e règle, qui découle en partie de la première, est recommandée. Si on raisonne dans le cadre d'un IRIS standard de 2 000 habitants, elle consiste, pour les variables qui concernent la population, à se limiter si possible à des croisements comportant dix modalités au plus. Même si, bien entendu, il y a peu de chances que toutes les modalités du croisement aient le même effectif, cela permet d'assurer, en moyenne, un effectif par case suffisant.

Concrètement, on peut imaginer :

- de décliner une variable en dix modalités au plus ;
- de croiser une variable à deux modalités avec une variable à cinq modalités au plus ;
- de croiser une variable avec trois modalités avec une variable ayant aussi trois modalités au plus.

IV - Que peut-on faire au niveau de l'IRIS ?

A partir de ces deux règles, on peut déterminer de « grands indicateurs » ou des tableaux pour lesquels la fiabilité de l'estimation est assurée. Pour ce faire, on distingue les variables relatives à la population de celles relatives au logement.

A - Population

La **population totale** est le premier des grands indicateurs. Ensuite, la pyramide des âges en 5 tranches d'âge croisées avec le sexe (soit 10 modalités au total) doit vérifier la 2e règle. Prenons comme exemple un IRIS de la ville d'Angers (*IRIS 0902 « Lareveillere »*) :

	Ensemble	Hommes	Femmes
De 0 à 19 ans	426 (9,7% - 8,7%)	228	198
De 20 à 39 ans	618 (5,6% - 7,2%)	300	318
De 40 à 59 ans	469 (7,1% - 8,3%)	221	248
De 60 à 74 ans	267 (6,6% - 11,0%)	92	175
75 ans ou plus	187 (7,9% - 13,2%)	64	123
	1967 (3,4% - 4,1%)	905	1062

Le 1^{er} pourcentage entre parenthèses donne le CV obtenu avec les simulations pour la commune d'Angers tandis que le 2^e pourcentage donne le CV avec la formule approchée de la partie I.

On voit que le CV donné par la formule a plutôt tendance à surestimer le coefficient de variation¹²⁸. Cela est dû au fait que les tranches d'âge retenues dans le tableau sont celles des variables d'équilibrage. Bien entendu, il est possible de s'intéresser à d'autres tranches d'âge du moment que leur effectif est suffisant.

Un autre domaine très regardé est celui de la **population active** et notamment du nombre de chômeurs. Toujours pour le même IRIS d'Angers, on pourrait obtenir le tableau suivant.

¹²⁸ La tranche 0-19 ans offre un contre-exemple. L'explication est la suivante : lors du tirage de l'échantillon annuel des 8% de logements, les variables d'équilibrage ne s'appliquent pas au même niveau géographique :

- les 5 tranches d'âge (0-19 ; 20-39 ; 40-59 ; 60-74 ; 75 ou +) sont assurées au niveau communal ;
- le nombre de logements (individuels/collectifs) est assuré au niveau de l'IRIS.

Pour les tranches d'âge, le niveau retenu est donc la commune. L'IRIS en profite indirectement mais pas aussi systématiquement que si l'équilibrage avait eu lieu à l'IRIS.

Cela rejoint l'idée évoquée au début de l'annexe, selon laquelle, à effectif égal, par exemple 200 personnes dans un IRIS ou 200 dans une commune, la précision sera meilleure pour le chiffre communal que pour le chiffre à l'IRIS.

â g e	Ensemble			Hommes			Femmes		
	Population active	Ayant un emploi	Chômeurs	Population active	Ayant un emploi	Chômeurs	Population active	Ayant un emploi	Chômeurs
15-19	18	16	2	14	12	2	4	4	0
20-24	97	73	24	43	32	11	54	41	13
25-29	211	180	31	107	92	15	104	88	16
30-34	120	108	12	63	60	3	57	48	9
35-39	100	89	11	52	48	4	48	41	7
40-44	131	114	17	63	55	8	68	59	9
45-49	123	115	8	57	53	4	66	62	4
50-54	112	96	16	66	59	7	46	37	9
55-59	41	35	6	20	17	3	21	18	3
60-64	6	5	1	2	2	0	4	3	1
65 ou +	1	1	0	1	1	0	0	0	0
	960	832	128	488	431	57	472	401	71

Ce tableau pourrait même être assorti d'un taux de chômage par sexe et tranche d'âge pour l'IRIS... Mais :

- il y a trop de cases non significatives du fait de leur faible effectif. Indépendamment de leur insuffisante fiabilité, ces chiffres risquent de varier fortement d'une année sur l'autre, sans que cela traduise en quoi que ce soit une évolution réelle ;
- plus profondément, y a-t-il une quelconque utilité à savoir qu'il y a une femme chômeuse ayant entre 60 et 64 ans dans l'IRIS 0902 d'Angers ?¹²⁹
- les données (de même que celles des anciens recensements généraux) sont redressées pour corriger la non-réponse. Cela signifie qu'à un effectif de 19 personnes dans un tableau issu du recensement de la population de 1999 pouvaient fort bien correspondre dans la réalité 22 personnes par exemple.

Il est clair que la production de tels tableaux est techniquement possible mais que leur publication est à déconseiller fortement. Dès lors qu'il n'est plus pertinent de faire des tableaux « comptables », comment regrouper les cellules du tableau précédent pour qu'à la fois il ait un intérêt pour l'utilisateur et que la précision soit suffisante ?

On est naturellement amené à regrouper les tranches d'âge qui sont trop détaillées. Pour ce regroupement, on peut reprendre les tranches le plus souvent retenues dans les fiches-profil « Activité : formes d'emploi » au niveau de l'IRIS :

¹²⁹ On néglige ici les problèmes relatifs aux risques d'identification de la personne, qui sont hors sujet mais évidemment pas à ignorer.

	Ensemble			Hommes			Femmes		
	Population active	Ayant un emploi	Chômeurs	Population active	Ayant un emploi	Chômeurs	Population active	Ayant un emploi	Chômeurs
15-24	115	89	26	57	44	13	58	45	13
25-49	685	606	79	342	308	34	343	298	45
50 ou +	160	137	23	89	79	10	71	58	13
	960	832	128	488	431	57	472	401	71

Les effectifs pour la population active et la population ayant un emploi sont désormais plus importants pour chaque tranche d'âge considérée. Par contre, la décomposition par sexe et tranche d'âge du nombre de chômeurs ne permet pas d'assurer a priori une qualité suffisante des résultats si l'on se réfère aux règles de précaution précédentes. Limiter l'information au nombre total de chômeurs (128) peut paraître insuffisant pour certaines comparaisons territoriales. Il peut paraître déroutant que les résultats relatifs à la population active et à la population ayant un emploi soient fiables, mais non ceux qui s'en déduisent comptablement. Cela tient, naturellement, à la faiblesse des effectifs considérés. De la même manière, il est fréquent que des estimations statistiques de soldes soient moins précises que celles de leurs composantes.

Cette difficulté est levée en revenant aux questions auxquelles on souhaite apporter des éléments de réponse et en cherchant les indicateurs statistiques les plus pertinents. En clair, les questions structurantes que l'on se pose sur un IRIS peuvent être les suivantes :

- quel est le nombre de chômeurs de l'IRIS ?
- quel est l'incidence de l'âge sur cet IRIS ?
- l'exposition au chômage des hommes et des femmes est-elle la même ?

Avec les données des recensements généraux de la population, la réponse à ces questions passait par des tabulations et une mise en forme d'effectifs. Avec les données du nouveau recensement, l'outil doit être adapté à l'objectif. Les indicateurs nécessaires pour répondre aux questions précédentes pourraient être les suivants :

- le nombre estimé de chômeurs de l'IRIS ;
- l'âge moyen des chômeurs de l'IRIS ;
- la proportion respective d'hommes et de femmes au chômage dans l'IRIS.

La demande relative à l'emploi sur l'IRIS d'Angers considéré pourrait donner lieu, dans ces conditions, à la production des informations statistiques suivantes :

- un tableau de cadrage :

	Ensemble		Hommes		Femmes	
	Popula- tion active	Ayant un emploi	Popula- tion active	Ayant un emploi	Popula- tion active	Ayant un emploi
15-24	115	89	57	44	58	45
25-49	685	606	342	308	343	298
50 ou +	160	137	89	79	71	58
	960	832	488	431	472	401

- un tableau de résultats sur le chômage :

Nombre de chômeurs	Age moyen des chômeurs (hommes)	Age moyen des chômeurs (femme)	Proportion de femmes chômeuses
128	34,3 ans	34,7 ans	56%

B - Logements

Outre des données sur les personnes, le recensement fournit des données sur les logements. La démarche de la partie A sur la population s'applique aussi aux logements.

La catégorie du logement (résidence principale, logement vacant, etc.) ainsi que le nombre de pièces d'habitation font partie des variables de base pour l'étude du logement. Le tableau relatif au même IRIS d'Angers, pour ces deux variables croisées, est le suivant :

	Nombre de pièces d'habitation						Nombre moyen de pièces
	1	2	3	4	5	6 ou +	
Résidences principales	70	169	232	302	130	48	3,42
Logements occasionnels	6	1	0	1	0	0	1,50
Résidences secondaires	0	0	2	0	0	0	3
Logements vacants	1	11	22	13	4	0	3,16
	77	181	256	316	134	48	3,39

Ce tableau ne respecte pas la 2e règle car il comporte 24 cases au lieu des 10 préconisées.

Après aménagement, la demande relative à la catégorie de logements et au nombre de pièces pourrait donner lieu à la production des informations statistiques suivantes :

- des données de cadrage :

Catégorie de logement	
Résidences principales	951
Logements vacants	51
Autres catégories de logements	10
	1012

Malgré un regroupement de cases, l'effectif « Autres catégories de logement » demeure faible. Il faut donc rester prudent sur l'interprétation.

	Nombre de pièces d'habitation					
	1	2	3	4	5	6 ou +
Ensemble des logements de l'IRIS	77	181	256	316	134	48

- des indicateurs :

	Nombre moyen de pièces		
Taux de vacance	Résidences principales	Logements vacants	Autres catégories de logements
5 %	3,42	3,16	1,8

C - Conclusion

Il ressort de ce qui précède qu'il apparaît nécessaire de proposer deux outils :

- un premier, défini en concertation avec les utilisateurs, reprenant les quelques tableaux en niveaux, structurants pour l'analyse des IRIS et fournissant de nouveaux indicateurs (taux, moyenne, médiane...) qui permettent d'avoir quand même des informations pertinentes ;
- un second permettant à l'utilisateur de faire ses regroupements mais l'alertant si besoin est sur la significativité des résultats obtenus.